Comparisons of Propensity Score Methods for Time to Event Outcomes: Evaluation through Simulations and Oral Squamous Cell Carcinoma Case Study

Sophie Yunfei Ma[1], Badr Id Said[2], Ali Hosni[2], Wei Xu[1,3], Sareh Keshavarzi[3]

[1]Biostatistics Division, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada
[2]Department of Radiation Oncology, Princess Margaret Cancer Centre, University of Toronto, Ontario, Canada
[3]Department of Biostatistics, Princess Margaret Cancer Centre, Toronto, Ontario, Canada.

Introduction

Backgrounds

- Observational studies are sometimes preferred over randomized clinical trials because of lower cost and more generalizability.
- It is recommended to use propensity score (PS) methods or covariate adjustment for confounding effect control.
- Few guidelines are available regarding the choice of PS approaches or covariate adjustment for the best performance in a given data.

Objectives

- In this study, we conducted Monte Carlo simulations comparing conventional covariate adjustment and eight common PS methods to estimate the average treatment effect for the overall population on time-to-event outcomes.
- We also applied the PS approaches to compare the effect of treatment and intervention on different time-to-event clinical outcomes in Oral Squamous Cell Carcinoma Cancer (OSCC) patients.

Methods

Monte Carlo Simulations

We examined 96 scenarios with the following factors: the prevalence of treatment, the event rate, sample size, and the true average treatment effect size (HR). Simulation was repeated for 1000 times for each scenario. Nine models were compared:

- 1:1 nearest neighbor matching (NNM) with or without caliper
- Stratification (5 or 10 strata)
- Inverse probability of treatment weighting (IPTW) with or without trimming
- Conventional covariate adjustment

Model Performance

- **Accuracy**: $\mathrm{E}(\hat{\beta}) = \frac{1}{1000}\sum_{j=1}^{1000}\hat{\beta}_j$, where $\hat{\beta}_j$ denotes the est. HR
- **Precision**: $2 * 1.96 * \sqrt{var(\hat{\beta}_j)}$, where $var(\hat{\beta}_j) = \frac{1}{1000}\sum_{j=1}^{1000}(\hat{\beta}_j - \mathrm{E}(\hat{\beta}))^2$
- **Mean squared error (MSE):** $E(B)^2 + var(\hat{\beta}_j)$,
  where the mean of bias $E(B) = \mathrm{E}(\hat{\beta}) - \beta_T$, where β_T denotes the true HR

Case Study

- **Population**: Adult patients with OSCC and have had the primary surgery.
- **Study design**: A single-institution retrospective study (2006-2018).
- **Sample**: 288 patients with tumor samples.
- **Treatment of interest**: if the patient had post-operative radiation therapy (PORT)
- **Intervention of interest**: if the patient had an engraftable tumor (Engraftment)
- **Outcome of interest**:
  - disease-free survival (DFS)
  - local-regional failure (LRF)
  - recurrence-free survival (RFS)
  - frequency and event rates are given in **Table 1**
- **Risk factors and confounders** include lymph node metastases, tumor size, and cancer grade.

Sample Size	n=288		
Disease-free Survival (DFS)	first recurrence or death 159 (55%)	last follow-up 129 (45%)	
Loco-regional failure (LRF)	first local or regional recurrence 99 (34%)	Last follow-up 133 (46%)	Deaths without failure 56 (19%)
Recurrence-free survival (RFS)	first local, regional, or distant recurrence 120 (42%)	Last follow-up 129 (45%)	Deaths without failure 39 (14%)

Table 1. Frequency and event rate of the three time-to-event outcomes

Results

Monte Carlo Simulations

NNM with caliper provided less precise estimates especially with few outcome events and low treatment prevalence. IPTW and covariate adjustment performed well in most cases and produced unbiased estimates with small uncertainty. (**Figure 1**)

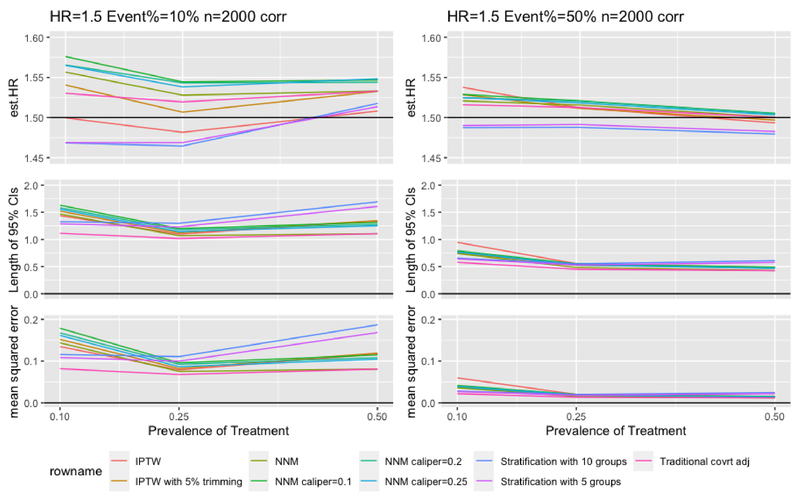

Figure 1. Comparison of eight PS methods and covariate adjustment with correlated covariates, n=2000. 1st row: estimated hazard ratio; 2nd row: mean length of 95% confidence interval; last row: mean squared error.

Case Study

- Standardized mean difference (SMD) < $\pm$0.25 (**Figure 2**)
- Results showed consistency across stratification and covariate adjustment (**Figure 3**)

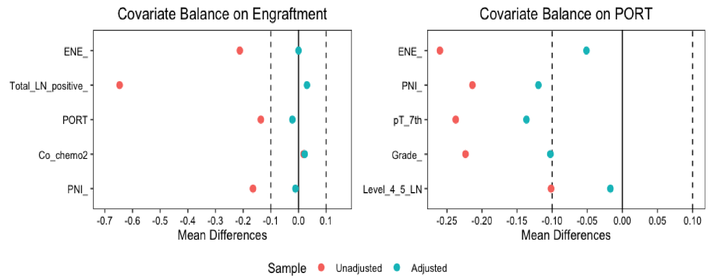

Figure 2. SMD before (red) and after (blue) NNM for Engraftment (left) and PORT (right); **ENE**, Pathologic extranodal extension; **Co_chemo2**, Concurrent chemotherapy; **PNI**, Perineural invasion; **pT_7th**, T-category primary tumor; **Grade**, Histological grade; **Level_4_5_LN**, Level IV-V nodal involvement.

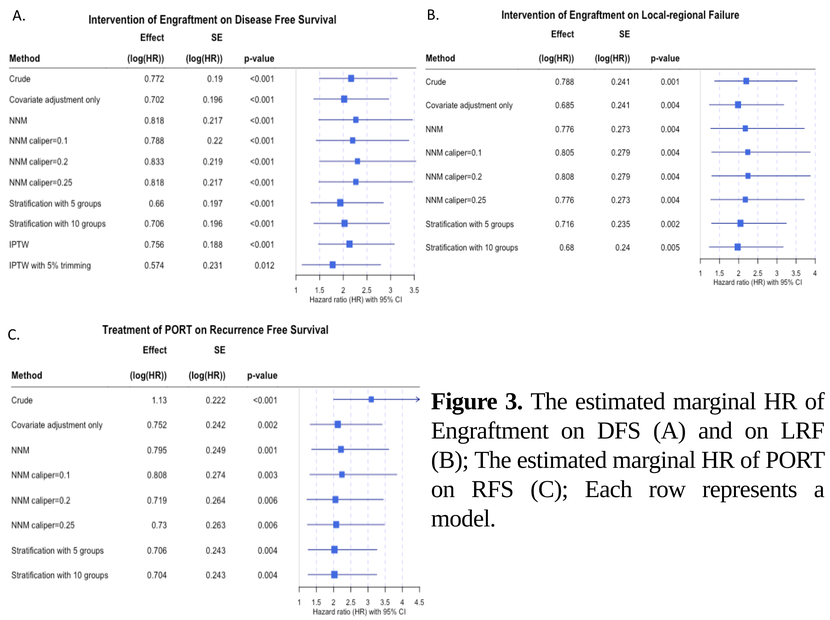

Figure 3. The estimated marginal HR of Engraftment on DFS (A) and on LRF (B); The estimated marginal HR of PORT on RFS (C); Each row represents a model.

Conclusion

Covariate adjustment and the IPTW method performed well across simulations and the case study. In practice, care should be taken to select the most suitable method when estimating the treatment, exposure or intervention effect on time-to-event outcomes.