WinKM: Approximating win-loss probabilities

Using overall and event-free survival functions

Lu Mao

Department of Biostatistics & Medical Informatics

University of Wisconsin-Madison

Visit https://lmaowisc.github.io/winKM

Outline

- **Introduction**
- **Methods**
  - Win-loss estimands
  - Approximating formula
  - `WinKM` workflow
- **Case studies**
  - A colon cancer trial
  - HF-ACTION trial
- **Conclusion**

Introduction

- **Win-loss statistics**
  - **Generalized pairwise comparisons** (GPC) (Buyse et al., 2025)
  - **Prioritization**: death > nonfatal event (hospitalization/disease progression)
  - **Proportions of wins vs losses**
    - Win ratio (Pocock et al., 2012), win odds (Brunner et al., 2021), net benefit (Buyse, 2010)
- **Meta analysis?**
  - Literature-wide evidence synthesis
  - Earlier studies not reporting win-loss measures
  - Patient-level data unavailable

Methods

Notation

- **Outcome data**
  - D_a: Overall survival (OS) time
  - T_a: Nonfatal event time
  - $T_a^* = D_a \wedge T_a$: Event-free survival (EFS) time
  - $a = 1$: treatment; 0: control
- **Summary functions**
  - **Joint survival:** $H_a(t, s) = \mathrm{pr}(T_a > t, D_a > s)$ - likely unavailable
  - **OS**: $S_a(t) = \mathrm{pr}(D_a > t)$ - available through Kaplan-Meier (KM) curve
  - **EFS**: $S_a^*(t) = \mathrm{pr}(T_a^* > t)$ - available through KM curve

Win-Loss Estimands

- **Win-loss probabilities** (Oakes, 2016)

$$
\begin{aligned}
w_{a,1-a}(\tau) &= \mathrm{pr}(\text{Group } a \text{ wins by time } \tau)\\
&= \mathrm{pr}\underbrace{(D_{1-a} < D_a \wedge \tau)}_{\text{Win on OS}} + \mathrm{pr}\underbrace{(D_1 \wedge D_0 > \tau, T_{1-a} < T_a \wedge \tau)}_{\text{Tie on OS, win on nonfatal}}\\
&= \int_0^\tau S_a(t-)\mathrm{d}F_{1-a}(t) \qquad\qquad \big(F_a(t) = 1 - S_a(t)\big)\\
&\quad + S_1(\tau)S_0(\tau)\int_0^\tau \mathrm{pr}(T_a > t \mid D_a > \tau)\mathrm{pr}(t \le T_{1-a} < t + \mathrm{d}t \mid D_{1-a} > \tau)
\end{aligned}
$$

  - **Win ratio:** $w_{1,0}(\tau)/w_{0,1}(\tau)$
  - **Win odds:** $w_{1,0}(\tau)/w_{0,1}(\tau)$
  - **Net benefit:** $w_{1,0}(\tau) - w_{0,1}(\tau)$

Survival-Conditional Event Rate

- **Second-term unknown**

$$\begin{aligned}
&\int_0^\tau \mathrm{pr}(T_a > t \mid D_a > \tau)\mathrm{pr}(t \leq T_{1-a} < t + \mathrm{d}t \mid D_{1-a} > \tau) \\
&= -\int_0^\tau H_a(t \mid \tau) H_{1-a}(\mathrm{d}t \mid \tau) \\
&= \mathrm{pr}(T_{1-a} < T_a \wedge \tau \mid D_1 > \tau, D_0 > \tau)
\end{aligned}$$

  - $H_a(t \mid \tau) = \mathrm{pr}(T_a > t \mid D_a > \tau)$: event-free probabilities in τ-survivors
  - *Association* between death and nonfatal event
  - Approximate it using $S_a(t)$, $S_a^*(t)$, and component-specific event counts

Approxmation - Idea

- **Start with t-survivors**

$$
\begin{aligned}
\mathrm{d}\Lambda_a(t \mid \tau) &:= \mathrm{pr}(t \leq T_a < t + \mathrm{d}t \mid T_a \geq t, D_a > \tau) \\
&\leq \mathrm{pr}(t \leq T_a < t + \mathrm{d}t \mid T_a \geq t, D_a \geq t) \\
&\approx \frac{N_a^{\mathrm{E}}}{N_a^*} \mathrm{d}\hat{\Lambda}_a^*(u)
\end{aligned}
$$

  - Notation
    - N_a^{E}: number of nonfatal events
    - N_a^*: number of composite (EFS) endpoints
    - $\hat{\Lambda}_a^*(t) = -\log \hat{S}_a^*(t)$: cumulative hazard of EFS
  - Inequality when *cross ratio* $\kappa(t, s) \geq 1$ (positive association)
  - Make up for bias by approximating $\kappa(t, s)$

Cross Ratio

- **Local dependence** (Oakes, 1982, 1986)

$$\kappa_a(t,s) = \frac{\mathrm{pr}(t \le T_a < t + \mathrm{d}t \mid T_a \ge t, D_a = s)}{\mathrm{pr}(t \le T_a < t + \mathrm{d}t \mid T_a \ge t, D_a \ge s)} = \frac{H_a(t,s)\partial^2 H_a(t,s)/(\partial t \partial s)}{\{\partial H_a(t,s)/\partial t\}\{\partial H_a(t,s)/\partial s\}}.$$

  - Relative change in nonfatal event risk at t with death at s

- **Under Gumbel–Hougaard copula** (Oakes, 1989)

$$\hat{\kappa}_a(t,s) \approx 1 + (\hat{\theta}_a - 1)\hat{\Lambda}^*_a(s)^{-1}$$

  - $\hat{\theta}$: estimated association parameter

Association Parameter

- **Estimating association parameter** (Mao et al., 2022)

$$\hat{\theta}_a = \frac{\log(1 - \hat{r}_a^{\mathrm{E}} / \hat{r}_a^*)}{\log(\hat{r}_a^{\mathrm{D}} / \hat{r}_a^*)} \vee 1$$

  - $\hat{r}_a^{\mathrm{E}} = N_a^{\mathrm{D}} / L_a^{\mathrm{OS}}$: nonfatal event rate
  - $\hat{r}_a^{\mathrm{D}} = N_a^{\mathrm{E}} / L_a^*$: death rate
  - $\hat{r}_a^* = N^* / L_a^*$: composite event rate
  - L_a^{OS}: total person-time at risk for OS
  - L_a^*: total person-time at risk for EFS

Approxmation - Formula

- **Formula**

$$H_a(t \mid \tau) \approx \prod_{0 \le u \le t} \left(1 - \frac{N_a^{\mathrm{E}}}{N_a^*} \mathrm{d}\hat{\Lambda}_a^*(u) \underbrace{\prod_{u \le s \le \tau} \left[1 - \{\hat{\kappa}_a(u, s) - 1\} \frac{\mathrm{d}\hat{F}_a(s)}{\hat{S}_a(s)} \right]}_{\text{Bias correction for } \tau\text{-survivorship}} \right)$$

- **Summary data needed**
  - $\hat{S}_a(t), \hat{S}_a^*(t)$: scan KM curves for OS and EFS (WebPlotDigitizer)
  - $N_a^{\mathrm{E}}, N_a^{\mathrm{D}}, N_a^*$: event counts reported in paper or CONSORT diagram
  - L_a^{OS}, L_a^*: total follow-up times calculated from risk table

`WinKM` Workflow

- A **step-by-step** approach
  - `prepare_km_data()`: read and clean digitized KM data
  - `merge_endpoints()`: align OS and PFS on a common time grid
  - `compute_increments()`: calculate $\mathrm{d}\hat{S}_a(t)$ and $\mathrm{d}\hat{S}^*_a(t)$
  - `compute_followup()`: derive total follow-up times from at-risk tables
  - `compute_theta()`: compute association parameters (θ_a)
  - `compute_win_loss()`: calculate final win/loss probabilities
- An **all-in-one** approach using `run_win_loss_workflow()`
- Visit package website for details

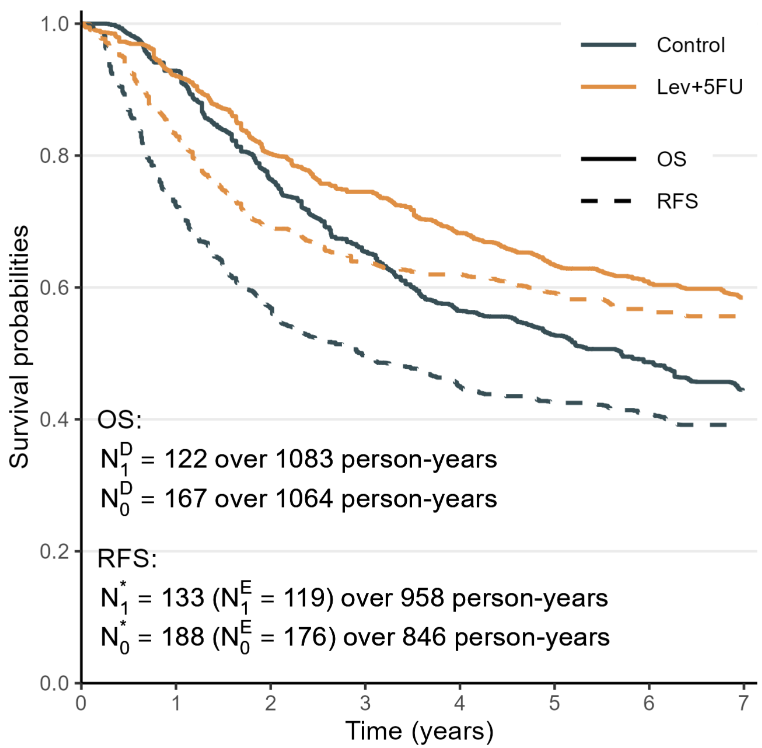
Case Studies

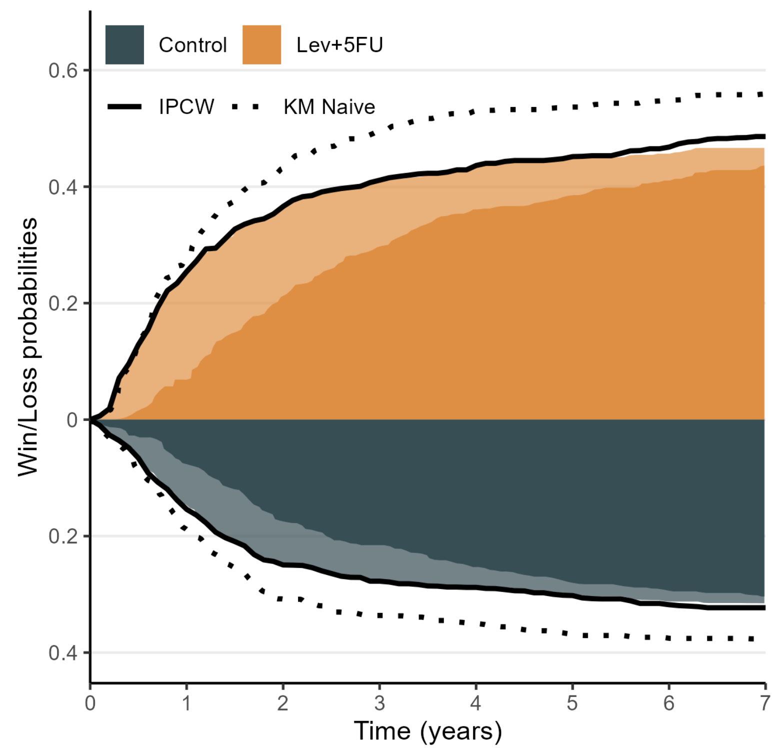
Colon Cancer Trial

- **Stage C disease** (Moertel et al., 1990)
  - Combined treatment (Lev+5FU; $n = 304$) vs control ($n = 314$ patients)
- **KM curves** for OS and relapse-free survival (RFS)
  - Extract estimates $\hat{S}_a(t)$ and $\hat{S}_a^*(t)$ from graphs using WebPlotDigitizer
  - Total event counts and person-years of follow-up from paper
- **Compare estimates of** $w_{1,0}(\tau)$ and $w_{0,1}(\tau)$
  - Adjusting for association between death and relapse (area plot)
  - KM naive: without adjustment
  - IPCW (reference): raw data-based estimates (Dong et al., 2020; Parner & Overgaard, 2024)

Colon Cancer Trial: Results

- **Left**: summary data; **right**: approximation results
  - 1-, 2-, 4-, and 7-year win ratios by the adjusted method (comparing with IPCW): 1.71 (1.65), 1.47 (1.47), 1.53 (1.51), and 1.48 (1.51), respectively



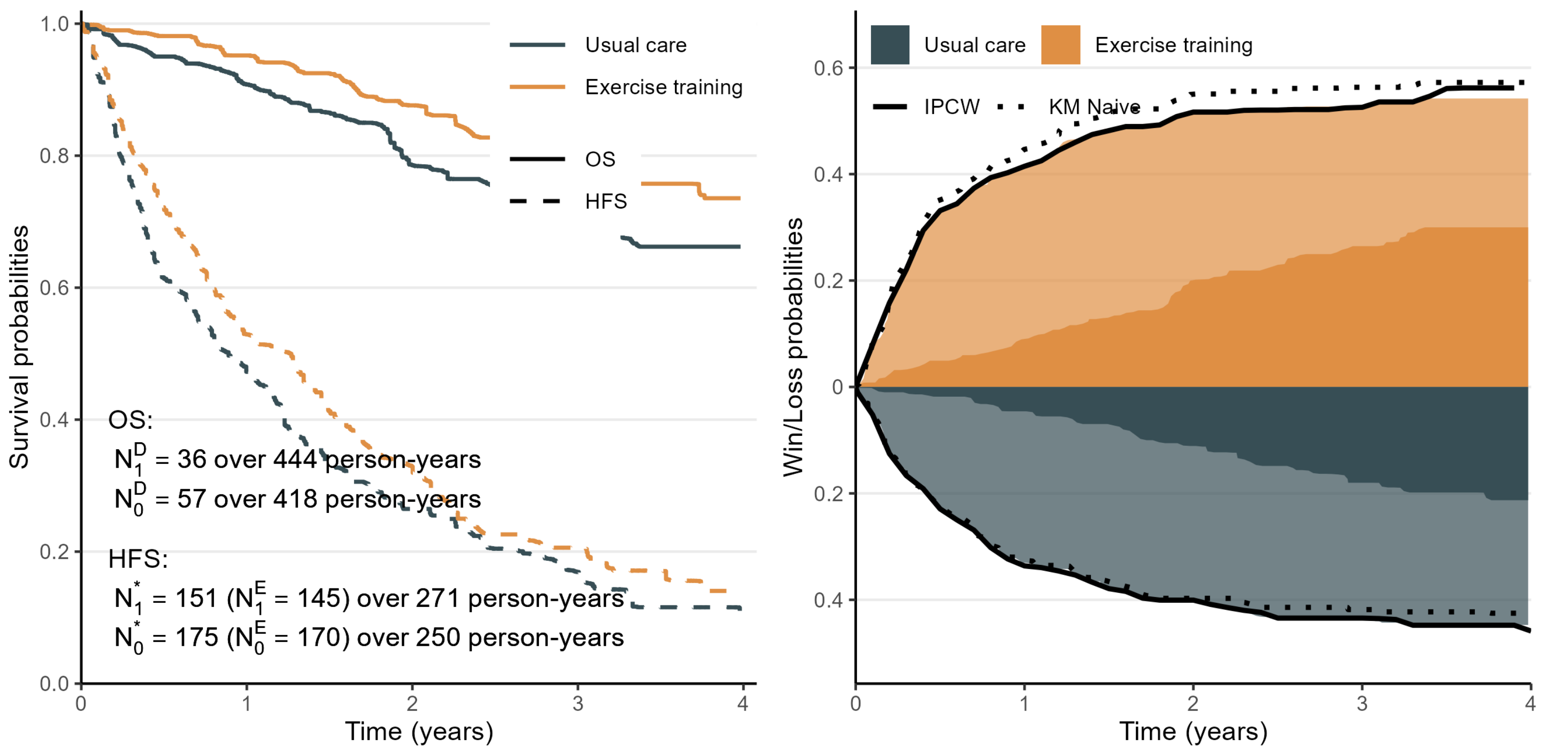


HF-ACTION Trial

- **Study background** (O'Connor et al., 2009)
  - 2,000+ heart failure patients from North America and France
  - Non-ischemic patients with baseline cardiopulmonary exercise test lasting 9 minutes or less (426 patients)
  - Exercise training ($n = 205$) vs usual care ($n = 221$)
- **KM curves** for OS and hospitalization-free survival (HFS)
  - Extract estimates $\hat{S}_a(t)$ and $\hat{S}^*_a(t)$ from graphs using WebPlotDigitizer
  - Total event counts and person-years of follow-up from paper

HF-ACTION Trial: Results

- **Left**: summary data; **right**: approximation results
  - 1-, 2-, 3-, and 4-year win ratios by the adjusted method (comparing with IPCW): 1.27 (1.23), 1.27 (1.29), 1.21 (1.21), and 1.21 (1.26), respectively



Conclusion

Additional Resources

- **Manuscript**

> Mao, L. (2025+) Approximating Win-Loss Probabilities Based on the Overall and Event-Free Survival Functions. Available at http://dx.doi.org/10.2139/ssrn.5142445

- **`WinKM` Website**

> Calculating win-loss statistics using summary data: https://lmaowisc.github.io/winKM/

Summary and Future Work

- `WinKM`: approximating win-loss measures using published OS and EFS data
  - Scanned KM curves
  - Event counts (deaths, EFS endpoints, nonfatal events)
  - Total follow-up times (from risk tables)
- **Adding standard errors**
  - Optimal combination of study-specific effect sizes
  - Assessment of between-study heterogeneity

References

Brunner, E., Vandemeulebroecke, M., & Mütze, T. (2021). Win odds: An adaptation of the win ratio to include ties. *Statistics in Medicine*, *40*(14), 3367–3384.

Buyse, M. (2010). Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. *Statistics in Medicine*, *29*(30), 3245–3257.

Buyse, M., Verbeeck, J., Saad, E. D., Backer, M. D., Deltuvaite-Thomas, V., & Molenberghs, G. (2025). *Handbook of Generalized Pairwise Comparisons*. Chapman; Hall/CRC. https://doi.org/10.1201/9781003390855

Dong, G., Mao, L., Huang, B., Gamalo-Siebers, M., Wang, J., Yu, G., & Hoaglin, D. C. (2020). The inverse-probability-of-censoring weighting (IPCW) adjusted win ratio statistic: An unbiased estimator in the presence of independent censoring. *Journal of Biopharmaceutical Statistics*, *30*(5), 882–899.

Mao, L., Kim, K., & Miao, X. (2022). Sample size formula for general win ratio analysis. *Biometrics*, *78*(3), 1257–1268.

Moertel, C. G., Fleming, T. R., Macdonald, J. S., Haller, D. G., Laurie, J. A., Goodman, P. J., Ungerleider, J. S., Emerson, W. A., Tormey, D. C., Glick, J. H., et al. (1990). Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. *New England Journal of Medicine*, *322*(6), 352–358.

O'Connor, C. M., Whellan, D. J., Lee, K. L., Keteyian, S. J., Cooper, L. S., Ellis, S. J., Leifer, E. S., Kraus, W. E., Kitzman, D. W., Blumenthal, J. A., et al. (2009). Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. *JAMA*, *301*(14), 1439–1450.

Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *44*(3), 414–422.

Oakes, D. (1986). Semiparametric inference in a model for association in bivariate survival data. *Biometrika*, *73*(2), 353–361.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, *84*(406), 487–493.

Oakes, D. (2016). On the win-ratio statistic in clinical trials with multiple types of event. *Biometrika*, *103*(3), 742–745.

Parner, E. T., & Overgaard, M. (2024). Estimation of win, loss probabilities, and win ratio based on right-censored event data. *Scandinavian Journal of Statistics*, *52*(1), 170–184.

Pocock, S. J., Ariti, C. A., Collier, T. J., & Wang, D. (2012). The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *European Heart Journal*, *33*(2), 176–182. https://doi.org/10.1093/eurheartj/ehr352