Predicting Pseudo Individual Patient-level Data with Covariates

Poster #MSR83

Verhoek A[1], Moradian H[2], Essink E[1], Heeg B[1]

Cytel, Inc., Rotterdam, Netherlands; [2]Cytel, Inc., Vancouver, Canada

Background

- All indirect treatment comparisons (ITC) rely on individual patient-level data (IPD) to compare the efficacy of healthcare interventions. However, IPD are often only available for the index study while Kaplan-Meier (KM) and aggregate data are used for the comparator trials. Techniques exist to generate pseudo-IPD from KM outcome data, but this only allows estimation of the outcome and not the covariates.

Objective

- This research aimed to develop a method to generate pseudo IPD by predicting the covariates on an individual patient level for use in ITCs.

Methods

- The "2-opt swap & shuffling" algorithm can be considered as an imputation problem to estimate an unknown covariate X_i for subsequent analysis using the information provided by a published paper. This method relies on:
  - KM data, (Y_i, S_i), time-to-event and event status for each treatment group, and related statistics (e.g., median survival time and number of events per covariate value per treatment)
  - Distribution of covariates (when x is categorical) for each treatment group
  - Coefficient estimate β and the related statistics (e.g., standard errors) of a reported Cox proportional hazards model:

$$h(t) = h_0(t) \times \exp(\alpha + \beta x + treatment).$$

- These inputs are used to calculate a normalized absolute error for any solution.
- A better solution can be found using the algorithm as visualized in Figure 1.

Experimental testing

- Several datasets were created to test the algorithm in practice:

  Scenario 1. Balanced covariate distribution and regular sample size: base-case scenario with the expectation that this would be the ideal scenario for the algorithm

  Scenario 2. Unbalanced covariate distribution: a skewed distribution between covariates (e.g., 30/70 split between male/female instead of 50/50) would give the algorithm both less and more space to optimize in the different subgroups.

  Scenario 3. Small sample size: Fewer patients reduces the solution space, making it easier for the algorithm to find the best solution.

  Scenario 4. Value of information—removing access to one or more pieces of information: Not all information could be available (e.g., the distribution of covariates might not be known). This scenario tests which information is critical.

- All scenarios (except the value of information) are conducted with three sub-scenarios:
  - a - Hazard ratios (HR) of both covariates are insignificant: insignificant HRs
  - b - One covariate is significant, the other is insignificant: semi-significant HRs
  - c - Both covariates are significant: significant HRs
- This gives a total of 10 scenarios.

Methods (cont.) and Results

Figure 1. "2-opt swap & shuffling" algorithm*

*Algorithm is here used colloquially, technically the "swap & shuffle" algorithm is a heuristic since no optimal solution is guaranteed.

Table 1. Overview of the experimental testing scenarios

Scenario	Hazard ratio	Male-female	Sample size
1a	Both insignificant	50%–50%	N=287
1b*	Semi-significant	50%–50%	N=287
1c	Both significant	50%–50%	N=287
2a	Both insignificant	30%–70%	N=287
2b	Semi-significant	30%–70%	N=287
2c	Both significant	30%–70%	N=287
3a	Both insignificant	50%–50%	N=140
3b	Semi-significant	50%–50%	N=140
3c	Both significant	50%–50%	N=140
4	Semi-significant	50%–50%	N=287

Figure 2. Visualization of one of the runs in scenario 1b (base case)

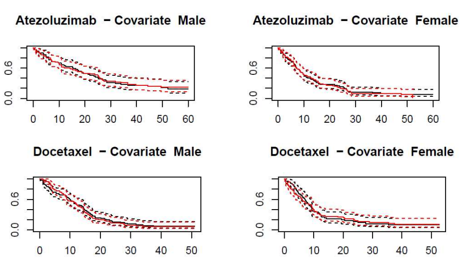

Table 2a. Results of scenario 1

Run	A - Insignificant	B - Semi-significant	C - Significant
1	1.696	0.682	1.492
2	2.247	0.261	0.466
3	1.576	2.087	1.387
4	0.314	0.317	1.095
5	1.595	1.653	2.297
Average	1.486	1.000	1.347
SD	0.710	0.825	0.664

Abbreviation: SD, standard deviation

Table 2b. Results of scenario 2

Run	A - Insignificant	B - Semi-significant	C - Significant
1	1.887	0.327	*Inf*
2	0.790	2.150	4.196
3	0.716	0.554	0.573
4	0.816	0.809	0.932
5	2.813	3.033	0.776
Average	1.404	1.374	NA
SD	0.924	1.167	NA

Abbreviations: inf, infinity (no solution found); NA, not applicable; SD, standard deviation

Table 2c. Results of scenario 3

Run	A - Insignificant	B - Semi-significant	C - Significant
1	0.257	3.459	0.507
2	0.593	0.329	0.709
3	0.742	4.191	3.107
4	0.354	2.268	1.561
5	0.286	0.775	1.184
Average	0.446	2.204	1.414
SD	0.211	1.665	1.032

Abbreviation: SD, standard deviation

Table 2d. Results of scenario 4*

Run	T-T-T	T-T-F	T-F-T	F-T-T	T-F-F	F-T-F	F-F-T	F-F-F
1	0.503	0.503	7.080	0.503	7.582	0.503	7.080	7.582
2	30.631	30.622	27.806	30.631	28.294	30.622	27.806	28.294
3	0.305	0.305	2.374	0.305	2.374	0.305	2.374	2.374
4	33.587	33.587	35.346	33.587	35.346	33.587	35.346	35.346
5	1.693	1.693	7.441	1.693	7.441	1.693	7.441	7.441
Average	13.344	13.342	16.010	13.344	16.208	13.342	16.010	16.208
SD	17.170	17.168	14.596	17.170	14.620	17.168	14.596	14.620

*The true-false is ordered per 1) number of events, 2) median survival, and 3) Brier score
Abbreviations: F, false; SD, standard deviation; T, true

Findings

- Scenario 1: The algorithm performed well with an almost perfect replication of the KM per subgroup.
- Scenario 2: Regenerating covariate values became more difficult, especially for the data set with only significant HRs. Here one run could not find a feasible solution
- Scenario 3: There was little effect in terms of mean error, but an increase in the standard error was observed.
- Scenario 4: The performance was mostly affected by: 1) no median and 2) combination of median and number of events.
- Sub-scenarios a/b/c: The semi-significant HR seemed to give better results, as there was more information to offset the covariates against. However, this was not necessarily the case in the other scenarios.

Conclusion

- This novel algorithm can recreate subgroup KM data for different covariates with minimal errors. The method has the potential for broader use in ITCs, such as effect-modifier testing and multivariate regression on IPD.