Poster #PRM232
Presented at the
ISPOR Europe 2018
November 10 to 14
BARCELONA, SPAIN

Longitudinal Matching: A Method for Generating Comparable Samples of Treated and Treatment-Naïve Patients with Progressive Conditions

Cook K.[1], Ali O.[1], Gupta D.[1], Ilin C.[2], Holmqvist D.[3], Lee D.[1], Tuttle E.[1], Bradt P.[4]

[1]Analysis Group, Menlo Park, CA, USA; [2]University of Wisconsin-Madison, Madison, WI, USA; [3]Vetcove LLC, New York, NY, USA; [4]Aegerion Pharmaceuticals, Cambridge, MA, USA

MOTIVATION

The effects of medical interventions on progressive conditions are ideally evaluated through randomized control trials (RCTs), in which patients are randomly assigned to a treatment or control group such that the two groups have comparable patient characteristics at study initiation.

However, it is often difficult, or even unethical, to conduct an RCT in order to evaluate the treatment of progressive conditions.

This leads to evaluations of novel treatments that are limited to using observational data.

As a result, regulatory bodies sometimes approve treatments in the absence of RCT-generated evidence:

- Hatswell et al. (2016) report that "Over the period [1999-2014], 76 unique indications were granted without RCT results (44 by the EMA and 60 by the FDA), demonstrating that a substantial number of treatments reach the market without undergoing an RCT."[1]

However, there are well-known limitations and biases associated with treatment effect estimation from observational data, highlighting the need for robust methods for estimating unbiased treatment effects.

We propose a method, Longitudinal Matching, that can be flexibly applied to observational data in order to approximate an RCT by matching treated patients with treatment-naïve patients at a similar stage of disease progression.

Unlike existing methods, Longitudinal Matching can be applied to data in which treated and treatment-naïve patients are observed at arbitrary times for different lengths of time.[2]

EXAMPLE

The following is a simple example of how the Longitudinal Matching algorithm matches each treated patient with a single untreated patient-history by choosing between eligible untreated patient-histories.

- There is one treated patient who is observed from age 15 onwards, represented by the orange line that starts at the circle in **Figure 1**. Suppose that this is also the time at which they initiate treatment.
- There are two untreated patients who are observed from birth until the end of data availability, represented by the blue lines in **Figure 1**. Crosses indicate the end of follow-up for all patients.
- Patients are matched on age and (integer-valued) stage of disease progression.
- Longitudinal matching compares the two untreated patients, at regular intervals, to the treated patient and attempts to find a point in their trajectories at which their characteristics most closely resemble those of the treated patient at the start of observation.
- Untreated patient 2 is chosen by the algorithm. Their data are left-truncated at age 12, and they are matched to the treated patient at treatment initiation (see **Figure 2**). The treated patient can now be compared to a counter-factual trajectory represented by untreated patient 2 from age 12 onwards.

Figure 1. Disease Progression of Treated Patient and Untreated Patients before Matching

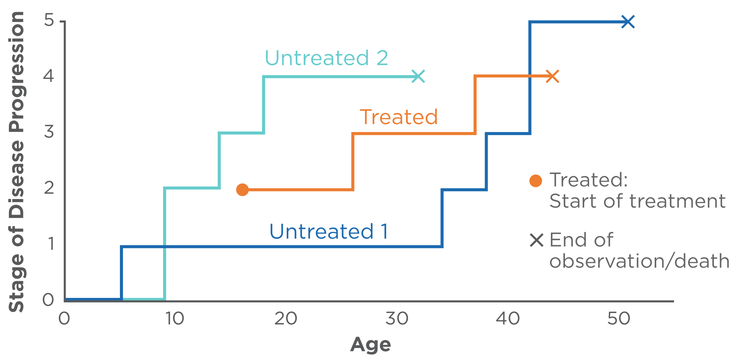
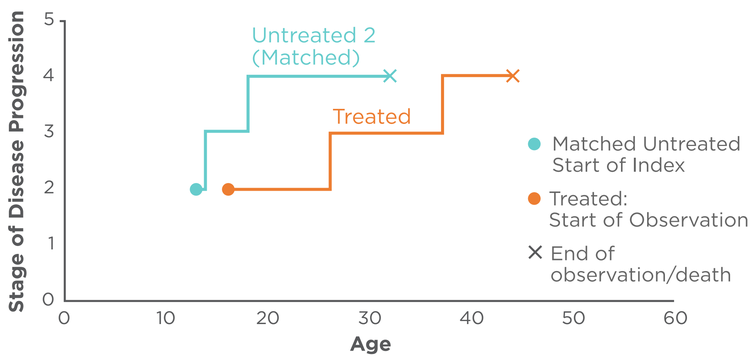


Figure 2. Disease Progression of Treated Patient and Untreated Patient 2 after Matching



METHOD: THE LONGITUDINAL MATCHING ALGORITHM

Suppose that observational data is available on both treated and untreated patients, and that the following conditions hold:

- Unconfoundedness — potential outcomes are independent of treatment assignment, conditional on covariates
- Monotonic progression — patients can only transition to more advanced stages of disease severity/impairment
- Treated patient data is left-truncated — it is available after treatment initiation, but possibly not before.

Denote treated patients by $i \in \{1,...,N\}$, and treatment-naïve (untreated) patients by $j \in \{1,...,M\}$.

Let the set of observations for patient i be $\{x_1^i,...,x_T^i\}$, where $t \in \{1,...,T\}$ denotes the time at which observations are recorded.

x_t^i can be a multidimensional vector with time-invariant entries (such as gender, state of residence, etc.) as well as time-varying entries, such as age, measures of disease progression, and mortality status.

For each treated patient, i, let t_i^* denote the time at which they received the treatment (without loss of generality, let $t_i^* = 1$ for all i).

The objective of the algorithm is to match each patient in the treatment group (at the time they receive treatment) to k treatment-naïve patients at the point in time when they are most similar.

Similarity is measured by the Mahalanobis distance between patients at a specific point in time across a group of chosen covariates.

Supposing we match on all available covariates, the Mahalanobis distance between treated patient i at time $t_i^* = 1$ and untreated patient j at time t (patient-history j_t), $m(i_1, j_t)$, is as follows:

$$m(i_1, j_t) = (x_1^i - x_t^j)^T \Sigma^{-1} (x_1^i - x_t^j)$$

Σ^{-1} represents the inverse of the variance-covariance matrix of matching variables in the appropriate sample.

METHOD (CONT'D)

When estimating the average treatment effect (ATE), the combined covariance matrix is used in the Mahalanobis distance metric, while the untreated group's covariance matrix is used instead when estimating the average treatment effect on the treated (ATT).[3]

The combined variance-covariance matrix is estimated as follows:

$$\Sigma_{ATE} = \left(\frac{N-1}{N+M-2}\right)\Sigma_T + \left(1 - \left(\frac{N-1}{N+M-2}\right)\right)\Sigma_C$$

$$\Sigma_{ATT} = \Sigma_C$$

In these covariance matrix expressions, Σ_T is the covariance matrix of the sample of treated patients, and Σ_C is the analogous matrix for the untreated patient-histories.

For each treated patient, i, the algorithm chooses the k untreated patient histories from the eligible set that minimize the distance $m(i_1, j_t)$. Denote the chosen set of patient-histories by $B(i;k)$ (for best k matches):

$$B(i;k) = \left\{j_t : j_t \in argmin(k)_w\{m(i_1, w)\}\right\}$$

In the above expression, $argmin(k)$ is an operator that finds the k smallest entries of Mahalanobis distance between treated patient i and untreated patient-histories, w.

The resulting matching pairs each treated patient, i, with k untreated patient-histories. Let M denote such a matching:

$$M = \{B(i;k)\}_{i=1}^N$$

When restrictions are placed on the number of times untreated patient-histories can be matched, the order in which treated patients are considered matters. Matchings can be expressed as functions of that order.

Let $I = \{1,...,N\}$ denote the set of treated patients. Let $P(I)$ denote a random permutation of I, $p^2(I)$ denote a random permutation of $P(I)$, and so on.

Let $S(M)$ denote the sum of Mahalanobis distances between each treated patient, i, and the k untreated patient-histories in $B(i;k)$ defined by matching M.

$$S(M) = \sum_{i=1}^N \left(\sum_{j \in B(i;k) \in M} m(i,j)\right)$$

The "best" matching, M*, is the one that minimizes $S(M)$ among a list of matchings generated by random permutations of I.

$$M^* = argmin\left\{S(M(I)), S(M(P(I))), ..., S(M(p^R(I))\right\}$$

The number of random permutations considered, R, is user-specified. The larger this number the better the matching M^* will be, but the longer it will take for the algorithm to complete its run.

APPLICATION

This section describes an application of Longitudinal Matching to estimate the treatment effect associated with metreleptin among patients with lipodystrophy.

Lipodystrophy is a progressive, heterogeneous disorder characterized by either a lack of or an abnormal distribution of adipose tissue.[4,5]

The condition can be characterized by adipose tissue loss to either specific areas (partial - PL) or the entire body (generalized - GL).[6]

Metreleptin is a drug that provides leptin replacement and is indicated as an adjunct to diet to treat the complications of leptin deficiency in patients with lipodystrophy.[6,7]

Data on 103 treated patients were obtained from a longitudinal, medical chart review studying the effects of metreleptin.[8]

Data on 230 untreated patients from 5 treatment centers were obtained from a retrospective, non-interventional, observational, closed cohort, longitudinal study assessing characteristics of lipodystrophy.[8]

Treated and untreated patient cohorts were not directly comparable due to differences in study design and patient characteristics.

- Untreated patients were part of a natural history study, in which they were observed from birth.
- Treated patients were selected into the clinical trial at a certain point in their life if they met disease severity criteria warranting treatment. Data collection began upon enrollment into the study.

Figure 3. Sample Characteristics of Treated and All Untreated Patients

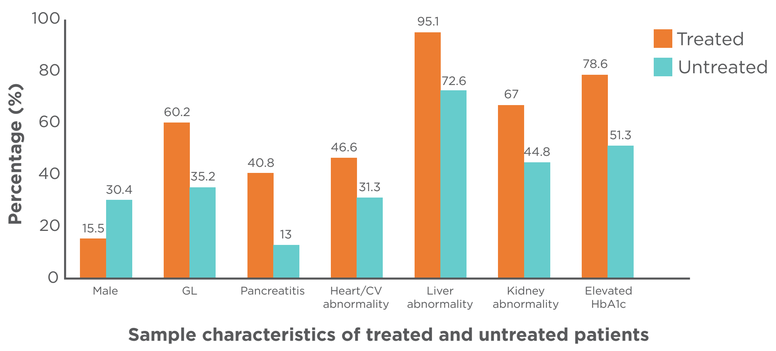


Figure 4. Sample Characteristics of Treated and Matched Untreated Patients

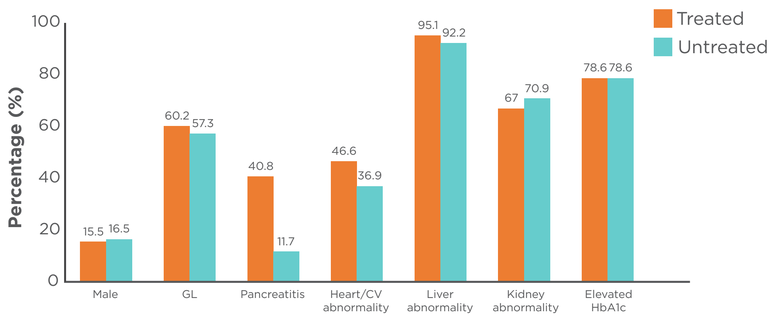


Figure 3 summarizes patient characteristics of the full set of treated and untreated patient cohorts, before matching.

- Treated patient characteristics are measured at baseline immediately before treatment initiation.
- Untreated patient characteristics are measured at the end of observation, when they are most impaired.

APPLICATION (CONT'D)

Treated patients are significantly farther along in disease progression than untreated patients, likely a result of trial inclusion restrictions.

Figure 4 summarizes patient characteristics of the full set of treated patients and the matched set of untreated patient-histories after implementing Longitudinal Matching on gender, GL/PL status, age at start of observation, the number of organ abnormalities, and an indicator of elevated HbA1c levels.

Untreated patient characteristics are, in general, no longer significantly different from the treated patients'.

SIMULATION

Longitudinal Matching was applied to data generated by simulating patient disease progression and mortality.

All patients begin in state 1 and face a Markov transition probability matrix governing their evolution through states 2,3,4,5 and death.

From\To	1	2	3	4	5	Death
1		0.0005				0.0002
2			0.0005			0.0004
3				0.0005		0.0006
4					0.0005	0.0008
5						0.0010

Later states correspond to later stages of disease progression, hence the probability of death is higher as the state increases.

We assume that treatment only affects mortality, and is simulated by a proportional decrease in all transition probabilities to death (for example, when the treatment effect is 0.5, the transition probability from state 4 to death is 0.0004 instead of 0.0008).

The entire trajectory of 2000 patients is simulated 100 times for up to 18250 periods. Treatment is then applied to half the patients at randomly chosen times in their histories, before the end of observation.

- Patient trajectories are re-simulated at these randomly chosen times using new (treatment-specific) mortality probabilities.

Untreated patients are matched on age and stage of disease progression (state). Longitudinal Matching was able to generate samples of almost perfectly balanced patients. For example, the mean index age across the 100 simulations with a treatment effect of 1 were 3.643 and 3.641, while the mean index states were 0.555 and 0.555 for treated and matched untreated patients, respectively. Note that matches generated from real world data are significantly less balanced.

The following table shows the mean and standard deviation (displayed in square brackets) of treatment effects across 100 simulations recovered by simple and full Cox proportional hazards models estimated on matched data (columns 3 and 4), and unmatched data (columns 5 and 6).

- The simple Cox model estimates the effect of treatment on mortality, while the full Cox model also controls for age and stage of progression at the index dates.

1	2	3	4	5	6
Number of Simulations	Actual treatment effect	Treatment effect (simple Cox model w/ matched data)	Treatment effect (full Cox model w/ matched data)	Naïve Treatment effect (simple Cox model w/ unmatched data)	Naïve Treatment effect (full Cox model w/ unmatched data)
100	0.25	0.274 [0.017]	0.256 [0.016]	0.316 [0.013]	0.225 [0.012]
100	0.50	0.521 [0.025]	0.497 [0.025]	0.607 [0.021]	0.449 [0.019]
100	0.75	0.753 [0.036]	0.735 [0.037]	0.883 [0.030]	0.671 [0.024]
100	1.00	0.970 [0.049]	0.970 [0.050]	1.150 [0.037]	0.887 [0.033]

Longitudinal Matching improved the estimation accuracy compared to a Cox model using unmatched data, recovering the true treatment effect.

CONCLUSION

Longitudinal Matching is a method that can be flexibly applied to rebalance longitudinal data across treated and untreated groups that are observed at arbitrary times.

It is particularly well-suited for the study of treatments for progressive conditions in the absence of an RCT.

An application to lipodystrophy estimates a protective treatment effect from data in which patients only qualified for the trial after meeting disease severity criteria exacerbating selection bias.

The validity of the method was verified through simulations of patient data from which Longitudinal Matching recovered the true treatment effects.

ACKNOWLEDGEMENTS

We would like to thank:

The clinical investigators who supported the collection of data regarding lipodystrophy:

- Baris Akinci and the clinical staff at Dokuz Eylul University, Izmir, Turkey.
- Rebecca J. Brown from the National Institute of Diabetes & Digestive & Kidney Diseases, and National Institutes of Health, Bethesda, MD, USA and her clinical team.
- Maria Cristina Foss de Freitas from University of Sao Paulo - Ribeirao Preto Medical School, Sao Paulo, Brazil and her clinical team.
- Renan Magalhaes Montenegro Jr. from Federal University of Ceara, Ceará, Brazil and his clinical team
- Elif A. Oral from University of Michigan, Ann Arbor, MI, USA and her clinical team.

Erica Birk, Andrea Di Miceli, Samantha Kaufhold, and Lamiya Sheikh from Analysis Group.

DISCLOSURES

This study was supported by Aegerion Pharmaceuticals. P. Bradt is an employee of Aegerion Pharmaceuticals.

K. Cook, O. Ali, D. Gupta, D. Lee, and E. Tuttle are employees of Analysis Group (AG). D. Holmqvist and C. Ilin are former employees of AG.

Analysis Group received consulting fees from Aegerion Pharmaceuticals.

REFERENCES

1. Hatswell AJ, Baio G, Berlin JA, Irs A, Freemantle N. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *BMJ* open. 2016;6(6):e011666-e011666.
2. Li YP, Propert KJ, Rosenbaum PR. Balanced risk set matching. *Journal of the American Statistical Association*. 2001;96(455):870-882.
3. Stuart EA. Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*. 2010;25(1):1.
4. Garg A. Acquired and inherited lipodystrophies. *New England Journal of Medicine*. 2004;350(12):1220-1234.
5. Huang-Doran I, Sleigh A, Rochford JJ, O'Rahilly S, Savage DB. Lipodystrophy: metabolic insights from a rare disorder. *The Journal of endocrinology*. 2010;207(3):245-255.
6. Brown RJ, Araujo-Vilar D, Cheung PT, et al. The Diagnosis and Management of Lipodystrophy Syndromes: A Multi-Society Practice Guideline. *The Journal of clinical endocrinology and metabolism*. 2016;101(12):4500-4511.
7. Myalepta [summary of product characteristics]. Amsterdam, The Netherlands: Aegerion Pharmaceuticals B.V.;2018.
8. Ali OA, Cook K, Gupta D, et al. Effect of Leptin Replacement Therapy (LRT) on Survival and Disease Progression in Generalized and Partial Lipodystrophy (GL, PL). *Diabetes.* 2018;67(Supplement 1).