ELSEVIER

Journal of Clinical Epidemiology 66 (2013) S99−S109

Journal of Clinical Epidemiology

Super learning to hedge against incorrect inference from arbitrary parametric assumptions in marginal structural modeling

Romain Neugebauer[a,*], Bruce Fireman[a], Jason A. Roy[b], Marsha A. Raebel[c], Gregory A. Nichols[d], Patrick J. O'Connor[e]

[a]*Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612, USA*
[b]*School of Medicine, University of Pennsylvania, PA, USA*
[c]*Institute for Health Research, Kaiser Permanente Colorado, CO, USA*
[d]*Center for Health Research, Kaiser Permanente Northwest, OR, USA*
[e]*HealthPartners Research Foundation, MN, USA*

Accepted 10 January 2013

Abstract

Objective: Clinical trials are unlikely to ever be launched for many comparative effectiveness research (CER) questions. Inferences from hypothetical randomized trials may however be emulated with marginal structural modeling (MSM) using observational data, but success in adjusting for time-dependent confounding and selection bias typically relies on parametric modeling assumptions. If these assumptions are violated, inferences from MSM may be inaccurate. In this article, we motivate the application of a data-adaptive estimation approach called super learning (SL) to avoid reliance on arbitrary parametric assumptions in CER.

Study Design and Setting: Using the electronic health records data from adults with new-onset type 2 diabetes, we implemented MSM with inverse probability weighting (IPW) estimation to evaluate the effect of three oral antidiabetic therapies on the worsening of glomerular filtration rate.

Results: Inferences from IPW estimation were noticeably sensitive to the parametric assumptions about the associations between both the exposure and censoring processes and the main suspected source of confounding, that is, time-dependent measurements of hemoglobin A1c. SL was successfully implemented to harness flexible confounding and selection bias adjustment from existing machine learning algorithms.

Conclusion: Erroneous IPW inference about clinical effectiveness because of arbitrary and incorrect modeling decisions may be avoided with SL.

Keywords: Super learning; Marginal structural model; Inverse probability weighting; Comparative effectiveness research; Time-dependent confounding; Selection bias

1. Introduction

In 2006, the American Diabetes Association changed its recommendations for the treatment of patients with type 2 diabetes mellitus (T2DM). The long-standing recommendation to begin pharmacotherapy only after a trial of lifestyle modification that failed to lower A1c to $<7\%$ was replaced with the new guideline for immediate prescription of metformin at detection of diabetes, regardless of A1c level. Authors of the new recommendation indicated that it reflects consensus rather than solid evidence.

In addition, adverse events linked to the use of thiazolidinediones [1] and inhaled insulin raised concerns over the long-term safety and effectiveness of agents used to control glycemia in T2DM patients. Although most experts interpret existing data as strongly supporting the safety and effectiveness of metformin, there is less confidence in the long-term safety and effectiveness of sulfonylurea and the use of metformin and sulfonylurea in combination.

Using the electronic health records (EHRs) from patients of four sites of the HMO Research Network (HMORN) Consortium [2], we assembled a large retrospective cohort study of adults with new-onset T2DM to evaluate the effect of immediate vs. delayed initial monotherapy or bitherapy with metformin and sulfonylurea on the risk of several clinical outcomes. We investigated these effects using marginal structural modeling (MSM) based on inverse probability

Conflict of interest/financial disclosure: G.A.N. received research support from AstraZeneca, GlaxoSmithKline, Merck & Co, and Takeda Pharmaceuticals America.

* Corresponding author. Tel.: +1-510-891-3234; fax: 510-891-3508.
E-mail address: romain.s.neugebauer@kp.org (R. Neugebauer).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.01.016

What is new?

Key findings

- Inferences from marginal structural modeling based on inverse probability weighting estimation and electronic health records data are sensitive to parametric decisions for modeling the treatment and right-censoring mechanisms.

What this adds to what was known?

- Super learning (SL) can successfully harness flexible confounding and selection bias adjustment from existing machine learning algorithms.

What is the implication and what should change now?

- Erroneous inference about clinical effectiveness because of arbitrary and incorrect parametric assumptions may be avoided with SL.

weighting (IPW) estimation for the purpose of properly accounting for the time-dependent confounding and informative selection bias that often arise in observational cohort analyses.

Here, our principal goal is twofold: (1) to illustrate the potential for incorrect inference resulting from inadequate parametric adjustment for confounding and informative censoring using MSM in comparative effectiveness research (CER) and (2) to illustrate the practical effect of and motivation for data-adaptive estimation with super learning (SL) in MSM. SL is a prediction algorithm, grounded in theoretical results, that builds an optimal weighted combination of predictors from a user-specified library of existing prediction methods using cross-validation. In addition, we illustrate the application of IPW estimation with a time-varying polychotomous (nonbinary) exposure. All illustrations are based on results for one survival outcome, the worsening of glomerular filtration rate (GFR).

2. An observational, multicenter, retrospective cohort study

We searched the entire adult membership of four participating HMORN health plans for enrollees meeting the eligibility criteria described in Appendix A. We enrolled each patient at the earliest date between January 1, 2006 and June 30, 2009, on which all criteria were met. As in a clinical trial, these eligibility criteria were devised to identify adults for whom the CER question is relevant, that is, adults with new-onset T2DM defined based on one elevated A1c measurement ($>6.5\%$) or two elevated measurements from fasting (>126 mg/dL) or random (>200 mg/dL) plasma glucose tests within a 2-year period. We excluded members whose life expectancy was limited by selected comorbid conditions. These criteria identified a cohort of 51,430 patients from which members with an observed or imputed baseline A1c $\geq 8\%$ were excluded. All 36,020 patients from the resulting cohort were followed up from the study entry until the earliest of June 30, 2010, plan disenrollment, or death.

3. Analytic approach

3.1. Motivation for MSM

To address the CER question, we aim to emulate inferences from an ideal randomized experiment with observational data [3]. In the hypothetical trial of interest, patients from the study cohort described previously would be randomized to one of several treatment arms corresponding with (1) no T2DM pharmacotherapy, (2) initiation of metformin monotherapy (met) at study entry, (3) initiation of sulfonylurea monotherapy (sul) at study entry, (4) initiation of bitherapy with metformin and sulfonylurea (met + sul) at study entry, (5) met initiation at 6 months after study entry, (6) sul initiation at 6 months after study entry, (7) met + sul initiation at 6 months after study entry, and (8) met initiation at 12 months after study entry, and so forth. This trial is ideal in the sense that patients would remain uncensored for the duration of the trial (2 years) and patients in arm (1) would comply with the assigned lack of therapy, whereas patients in all other arms would comply with the assigned treatment regimen until at least the assigned time of treatment initiation. In each arm, patients' GFR would be monitored to detect evidence of first GFR worsening after study entry. The corresponding survival curve in each arm would be contrasted at 2 years. More specifically, the cumulative risk differences between any two treatment interventions in this trial are the comparative effectiveness measures that we wish to evaluate with observational data.

Standard modeling approaches are known [4,5] to be inadequate to handle time-dependent confounding and selection bias [6] as they rely on conditioning of time-varying covariates, which are also often expected to lie on a causal pathway of interest between one of the variables defining the exposure groups of interest and the outcome. Fig. 1 illustrates such a scenario with a causal diagram [7,8] of a subset of measurements collected over 1 year for each patient in this study. MSM with IPW estimation can permit adequate adjustment for such time-varying covariates also on a causal pathway between early therapy exposure and the outcome and can directly emulate inference for the intention-to-treat (ITT) effects of interest [9−11] in this study (These effects are referred to as ITT effects because their interpretation is similar to the interpretation of