UMC Utrecht
Julius Center

On the aggregation of historical prognostic scores for causal inference

Thomas Debray, PhD

Julius Center for Health Sciences and Primary Care,
University Medical Center Utrecht, The Netherlands

Cochrane Netherlands, Utrecht, The Netherlands

thomasdebray.be

I have no actual or potential conflict of interest in relation to this presentation.
I would like to acknowledge Long Nguyen, corresponding author of the presented work.

Causal inference in non-randomized studies

Adjust for confounding bias

- Propensity Score Analysis (PPS)
  - Assess treatment effect among patients who have the same probability of receiving the treatment
- Prognostic Score Analysis (PGS)
  - Proposed by Hansen in 2008
  - Assess treatment effect among patients who have the same predicted prognosis for a given reference treatment

PPS and PGS assume absence of hidden bias, and therefore need to adjust for many confounders. This may be problematic when data are sparse.

Prognostic Score Analysis

"Traditional" methods for obtaining a prognostic score

- Develop from scratch using non-randomized data at hand
- Develop from scratch using a separate large sample of control individuals
- Use a previously published prognostic score

Alternative approach

- Aggregate multiple published prognostic scores, and tailor them to the control arm of the non-randomized treatment study
- This allows to adjust for a large number of confounders without having to re-estimate their individual effects
- Inspired by stacked regressions (Debray et al. 2014)

Prognostic Score Analysis

Aggregation of M prognostic scores $\hat{\Psi} = \left[\hat{\psi}_1, \hat{\psi}_2, \ldots, \hat{\psi}_M\right]$ is achieved by optimizing the following function in the control subjects from the non-randomized study:

$$argmin\,[Y \log(P) + (1 - Y) \log(1 - P)]$$

where

$$P = \mathrm{g}^{-1}\left(\theta_0 + \sum_m \theta_m g(\hat{\psi}_m(\mathbf{X}))\right)$$

with unknown parameters $\theta_0, \theta_1, \ldots, \theta_M$ and the constraint $\theta_m \geq 0$.

Case Study

Subgroup analysis in a clinical trial to compare the effectiveness of controlled asthma versus partly controlled asthma on the 1-year risk of exacerbation in asthmatic patients

- Total sample size: 281
- Total number of measured covariates: $p = 20$
- Estimation of average treatment effect in treated population (ATT)

	ATT	95% CI	p	dfr	EPV
Same-sample PGS	0.021	-0.054 to 0.097	7	8	3.3
Published PGS (Schatz)	0.018	-0.069 to 0.105	3	0	-
Published PGS (Eisner)	0.027	-0.054 to 0.108	2	0	-
Published PGS (TENOR)	-0.002	-0.088 to 0.085	10	0	-
Aggregated PGS	0.019	-0.060 to 0.098	15	4	7.7

EPV = Events per variable

Results from simulation study

- Development of PGS from scratch is recommended when datasets at hand are very large
- Published PGS should be updated to the control arm of the non-randomised study
  - to avoid bias due to miscalibration
  - to avoid defective adjustment of confounding bias
- Aggregating published PGS allows to adjust for many covariates
  - One unknown parameter for each PGS, rather than for each covariate
  - Potential to exclude PGS with poor fit in data at hand
- Aggregating published PGS enables unbiased estimation of treatment effects
  - even if existing scores ignore important covariates
  - even if existing scores estimate different (but related) prognostic endpoints