# Controling time-varying confounding in a cohort of bone marrow transplant patients, a comparison of g-methods.

Alexander Keil[a], Jessie K. Edwards [a], Ashley I. Naimi[b], Stephen R. Cole [a]

[a]Department of Epidemiology, University of North Carolina, Chapel Hill, NC; [b]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montréal

## Introduction

Graft-versus-Host-Disease (GvHD) a serious side-effect of bone marrow transplant (BMT), also helps kill residual leukemia cells

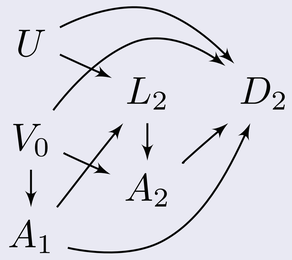
Time varying confounding for GvHD → mortality by platelet count and leukemia relapse (see figure 1)

G-methods can appropriately estimate effect of GvHD → mortality, choice between methods depends on causal knowledge

Without causal knowledge choice of methods is unclear

137 bone marrow transplant patients We estimate GvHD → mortality using 3 g-methods

## Causal graph




## Notation

$k=\{1,2,...\}$ subscript indexes time in days

$A_k$ GvHD at time k (yes or no)

$L_k$ Time varying confounders: relapse and platelet count (time varying confounders)

$V_0$ Baseline confounders: age (spline), sex, cytomegalovirus status (yes or no), leukemia type (AML vs. others), transplant wait time (linear)

$D_k$ Death between day k-1 and k

$T$ Time to all-cause mortality

## Methods

G-formula Using Monte Carlo approximation of parametric G-formula

- Pooled logistic model: generate conditional probabilities of all covariates
- Substitute probabilities under interventions
- Resample data under probabilities to generate data under interventions

Marginal structural mode Fit via IP weighting

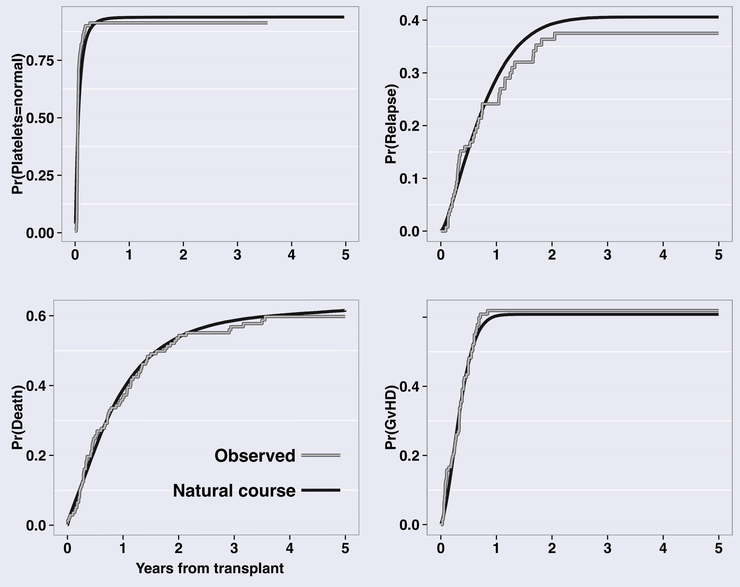
- Pooled logistic model: generate conditional probabilities of GvHD
- Inverse probability weight the data
- Fit a marginal model to re-weighted data

Structural nested accelerated failure time model Fit via g-estimation

- Propose a total effect of exposure (total effect=$\psi$)
- Generate $\boldsymbol{T^*}$ = residual outcome not due to exposure via $\boldsymbol{T^* = \int_0^T exp(GvHD_k \psi) dk}$
- Test $\boldsymbol{Pr(GvHD_k | prior\ confounders, T^*) = Pr(GvHD | prior\ confounders)}$
- Repeat 1-3 until exposure is not associated with "residual" outcome, that value of $\psi$ is the estimate

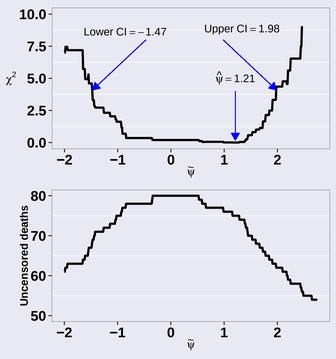
Comparing methods We generated HRs and 95% confidence intervals to compare 3 g-methods with standard Cox regression

## G-formula distributions



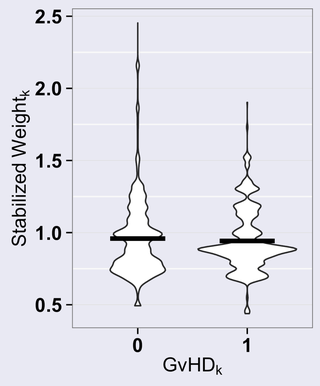

Distributions of time varying covariates: observed and G-formula under no intervention (natural course)

## SNAFTM g-function




The g-function (top panel) and number of uncensored deaths (bottom) as a function of the proposed total effect of GvHD on mortality

## MSM weights



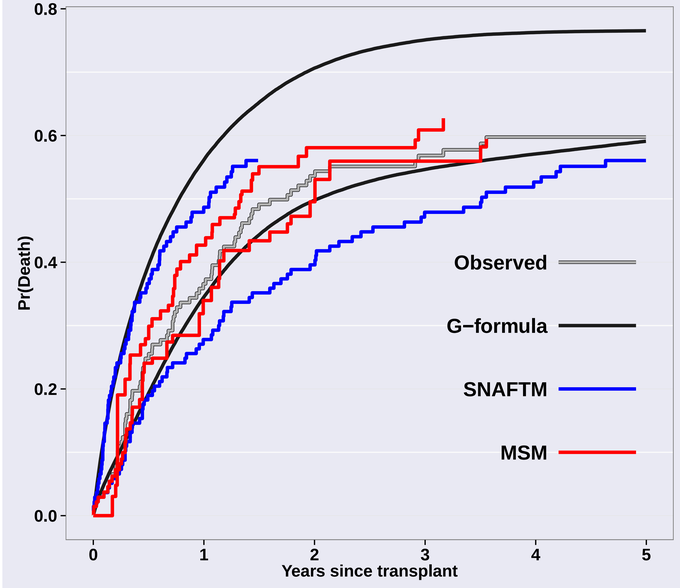

Distribution and mean of time specific weights used to fit marginal structural model.

## Hazard ratios

| | Cox HR | 95% CI |
|---|---|---|
| **Crude** | 1.24 | 0.77, 2.00 |
| **Baseline adjusted** | 1.16 | 0.66, 2.03 |
| **Fully adjusted** | 2.36 | 1.30, 4.29 |
| **G formula** | 1.83 | 0.98, 3.40 |
| **MSM** | 1.18 | 0.70, 1.98 |
| **SNAFTM** | 2.00 | |

Hazard ratios, 95% confidence intervals from conventional Cox regression models and 3 g-methods on Bone Marrow Transplant data.

## Survival curves




Survival curves generated by method by GvHD (always - top three lines vs. never = bottom three lines)

## Conclusions

- Inverse probability weighting may be biased due to positivity violations
- Structural nested models relax restrictions but at a cost of efficiency
- In the absence of causal knowledge the g-formula is useful as a comparator to less modeling intensive methods

## Funding sources

National Institute of Health Sciences Grant # T32 ES007018 (AK) NIH Grant #s R01CA117841 (JKE) and R01AI100654 (JKE and SRC) Post-Doctoral Research Award from the Fonds de recherche en santé du Québec (AIN)