Estimating the Causal Effect of Early Use of Erythropoietic Stimulating Agents in Intermediate-1 to Low-risk MDS Patients: An Application of the Longitudinal Targeted Maximum Likelihood Estimation

Yingying Zhang[1] Noemi Kreif[1] Alastair Bennett [1] Vijay Gc[2] Andrea Manca[1]

1. Centre for Health Economics, University of York, York, UK 2. School of Human and Health Sciences, University of Huddersfield, Huddersfield, UK

contact emails: yingying.zhang@york.ac.uk, andrea.manca@york.ac.uk

Introduction

- The paper uses the European Myelodysplastic Syndromes Registry and estimates average treatment effects (ATE) of using erythropoiesis-stimulating agents (ESA).
- We compare patients having initiated/received ESA for certain time periods with patients having never ever received ESA for the same amount of time.
- We account for time-varying confounding that affect both the treatment decisions and patient outcomes

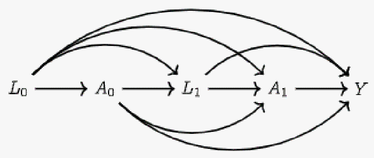
Longitudinal TMLE

- **Double Robust**
  only one of outcome (Q) or intervention mechanism (g) is needed for consistency of the final estimate
- **Machine Learning**
  super learner algorithm can be used to estimate components of Q and g while retaining valid statistical inference.
- **Uncertainty**
  SEs are correct asymptotically when estimates of Q and g are both consistent, and are conservative when only g is estimated consistently.
- **Various Outputs**
  mean outcomes for treatment and control groups; ATE, ATT, causal risk ratio/ causal odds ratio

Data and Variables

One observation O is coded as

$O = (L_0, A_0, C_0, Y_1, .., L_T, A_T, C_T, Y_{T+1})$, where



L nodes are covariates

A nodes are intervention nodes, 1 if being in the ESA group and 0 not

C nodes are censoring nodes, 1 if being uncensored and 0 not

Y nodes are outcomes including EQ-5D scores and cumulative survival probabilities.

Baseline covariates: Age at baseline, MDS co-morbidity index

time-varying confounders: haemoglobin levels

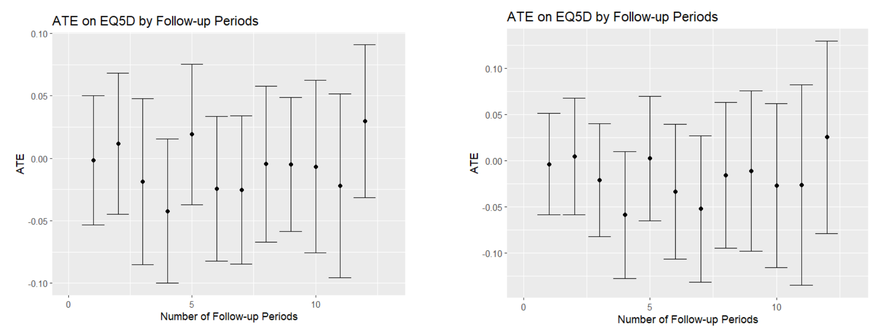
time-varying covariates: haemoglobin levels, bone marrow blasts greater than or equal to 5%, Karnofsky performance status, platelets count categories, absolute neutrophil count

Super Learner

- SL algorithm uses cross-validation to find the optimal weighted convex combination of multiple candidate prediction algorithms.
- Machine learning algorithms can improve the specifications of Q and f functions.
- The algorithms we use include generalized linear model, Stepwise regression, neural network, generalized additive models, Elastic net

ESA Initiation Analysis: ITT Effects

- A sample of 440 individuals for the QoL analysis, and a sample of 519 individuals for survival analysis
- We compare patients having initiated ESA for certain time periods with patients having never ever taken ESA for the same amount of time.
- We account for time-varying confounding that affect both the treatment decisions and patient outcomes



(1) with super learner (2) with GLM

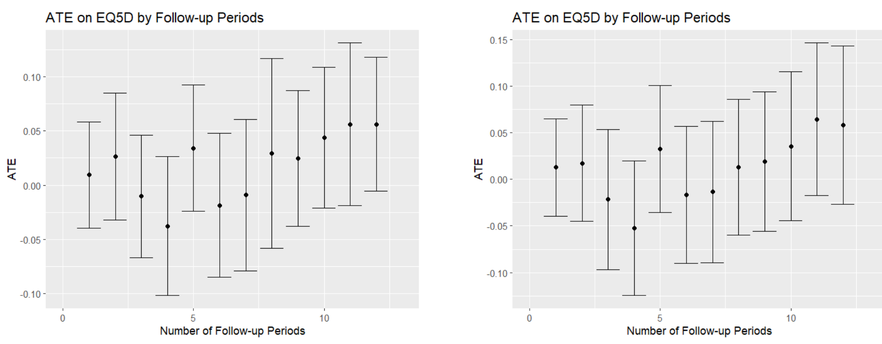
Figure 1: SL Estimations of ATEs of Initiating ESA for QoL Analysis

(1) with super learner (2) with GLM

Figure 2: SL Estimations of ATEs of Initiating ESA for Survival Analysis

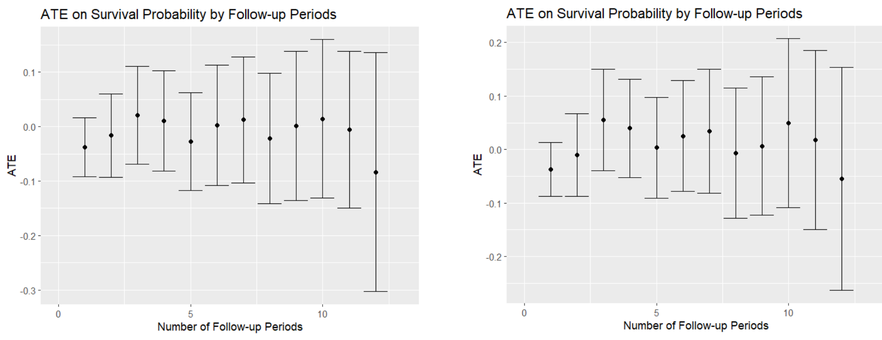
ESA Receipt Analysis: ATE Effects

- A sample of 425 individuals for the QoL analysis, and a sample of 519 individuals for survival analysis
- We compare patients having received ESA for certain time periods with patients having never ever taken ESA for the same amount of time.



(1) with super learner (2) with GLM

Figure 3: SL Estimations of ATEs of Receiving ESA for QoL Analysis



(1) with super learner (2) with GLM

Figure 4: SL Estimations of ATEs of Receiving ESA for Survival Analysis

Findings

- Using ESA has no significant effect on patients' EQ5D or on cumulative survival probabilities across all treatment time periods.
- Accounting for time-varying confounding reduces the bias of the predicted ATEs and intervention-specific mean outcomes.

HTx Next Generation Health Technology Assessment

CHE Centre For Health Economics