Targeted Maximum Likelihood Estimation (TMLE)

HPDS Causal Reading Club

Rebecca Hurwitz

May 20, 2025

Stanford University

Agenda

Schuler and Rose: TMLE for Causal Inference in Observational Studies

1. Why TMLE?
2. Motivating Example
3. Machine Learning Integration & Simulation Study
4. Takeaways

Torres et al.: US Migration Status of Adult Children and Cognitive Decline Among Older Parents

5. Background & Study Design
6. TMLE in Practice
7. Takeaways

American Journal of Epidemiology
© The Author 2016. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.
Vol. 185, No. 1
DOI: 10.1093/aje/kww165
Advance Access publication:
December 9, 2016

Practice of Epidemiology

Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies

Megan S. Schuler and Sherri Rose*

* Correspondence to Dr. Sherri Rose, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02215 (e-mail: rose@hcp.med.harvard.edu).

Stanford University

Why TMLE?

- Observational data are increasingly used for causal inference in epidemiology
- Traditional methods (G-computation, propensity scores) can fail under model misspecification
- **TMLE** is a doubly robust method designed to improve reliability
- This paper makes TMLE accessible and empirically tests its benefits

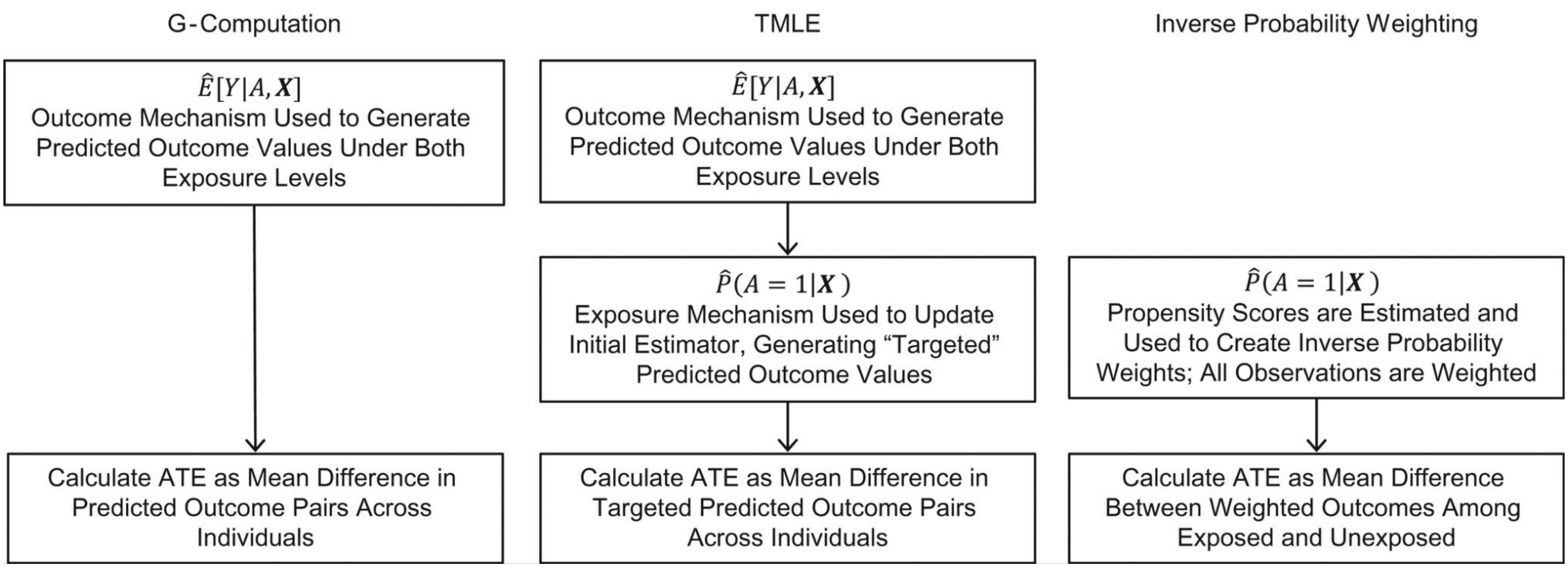
Causal Inference Framework Refresher

- Let's define:
  - *A:* binary exposure
  - *X*: vector of covariates
  - *Y*: continuous outcome
  - Y_1, Y_0: potential outcomes (pair in this paper)
- Average Treatment Effect (ATE):
  - ATE = $E[Y_1 - Y_0]$
- Assumptions we need to interpret statistical estimates causally:
  - SUTVA (Stable Unit Treatment Value Assumption)
  - No unmeasured confounding
  - Positivity: $0 < P(A = 1 \mid X) < 1$

TMLE Overview

- TMLE combines outcome and exposure models into one estimator
- Updates initial prediction to optimize bias-variance for parameter of interest
- Doubly robust: Only one model needs to be correctly specified
- Substitution estimator: can easily leverage machine learning

Estimators for ATE



Motivating Example

What is the causal effect of regular physical exercise ($\geq$ 150 minutes/week) on depressive symptoms?

Motivating Example: Simulated Data

- **Exposure:** Regular exercise (binary)
- **Outcome:** CES-D score (0-60); lower = fewer symptoms
- **Confounders:** sex, baseline psychosocial therapy for depression, baseline antidepressant use
- **True Data-Generating Mechanism:**
  - A more likely for women, those already in therapy or on meds
  - Y reduced more by exercise if also on antidepressants
- **True ATE:** −3.38 CES-D points

TMLE Steps

1. Generate initial estimate of E($Y \mid A , X$)

2. Estimate exposure mechanism P($A = 1 \mid X$)

3. Update initial estimate of E($Y \mid A , X$)

4. Generate targeted estimate of target parameter:

$$\widehat{ATE} = \frac{1}{n} \sum_{i=1}^{n} [\hat{Y}_1^* - \hat{Y}_0^*]$$

Machine Learning Integration & Simulation Study

Why use ML in the context of TMLE?

- True functional forms $E(Y \mid A, X)$ and $P(A = 1 \mid X)$ may be complex
- TMLE supports flexible, data-adaptive algorithms:
  - LASSO, random forest, regression trees, GAMs, etc.
- Optimal algorithm selection is challenging -→ Super Learner
- **Improves robustness by reducing model misspecification and boosting predictive accuracy**

Simulation Study Design

- Compare TMLE, G-computation, and IPW under:
  - Parametric misspecification (main terms or omitted variable)
  - Machine learning (super learning)

- **Setup:**
  - 1,000 replications each with 1,000 patients
  - Same data-generating model as motivating example
  - True ATE: −3.38 CES-D points

Simulation Study Design

- Misspecification Conditions:

Outcome model:

- Main terms only (omits interaction)
- Omitted variable (drops antidepressant use)

Exposure model:

- Omitted variable (drops antidepressant use)

Table 1. Estimates of the Mean Average Treatment Effect, Mean Bias, and 95% Confidence Interval in a Simulation Study of 3 Different Estimation Methods (Targeted Maximum Likelihood Estimation, G-Computation, and Inverse Probability Weighting)[a]

Estimator	Mean ATE (SE)	Mean Bias	95% CI
Targeted Maximum Likelihood Estimation			
Super learner			
Outcome variables: A, X_1, X_2, X_3; exposure variables: X_1, X_2, X_3	−3.39 (0.35)	−0.01	−4.05, −2.64
Misspecified parametric regression			
Main-terms misspecification			
Outcome variables: A, X_1, X_2, X_3	−3.39 (0.35)	−0.01	−4.08, −2.64
Omitted-variable misspecification			
Outcome variables: A, X_1, X_2	−3.39 (0.36)	−0.01	−4.09, −2.63
Exposure variables: X_1, X_2	−3.39 (0.35)	−0.01	−4.07, −2.69
G-Computation			
Super learner			
Outcome variables: A, X_1, X_2, X_3	−3.27 (0.35)	0.11	−3.98, −2.56
Misspecified parametric regression			
Main-terms misspecification			
Outcome variables: A, X_1, X_2, X_3	−3.25 (0.33)	0.13	−3.91, −2.59
Omitted-variable misspecification			
Outcome variables: A, X_1, X_2	−4.98 (0.37)	−1.60	−5.69, −4.24[b]
Inverse Probability Weighting			
Super learner			
Exposure variables: X_1, X_2, X_3	−3.43 (0.37)	−0.05	−4.17, −2.63
Misspecified parametric regression			
Omitted-variable misspecification			
Exposure variables: X_1, X_2	−4.96 (0.37)	−1.58	−5.67, −4.21[b]

Table 1. Estimates of the Mean Average Treatment Effect, Mean Bias, and 95% Confidence Interval in a Simulation Study of 3 Different Estimation Methods (Targeted Maximum Likelihood Estimation, G-Computation, and Inverse Probability Weighting)[a]

Estimator	Mean ATE (SE)	Mean Bias	95% CI
Targeted Maximum Likelihood Estimation			
Super learner			
Outcome variables: A, X_1, X_2, X_3; exposure variables: X_1, X_2, X_3	−3.39 (0.35)	−0.01	−4.05, −2.64
Misspecified parametric regression			
Main-terms misspecification			
Outcome variables: A, X_1, X_2, X_3	−3.39 (0.35)	−0.01	−4.08, −2.64
Omitted-variable misspecification			
Outcome variables: A, X_1, X_2	−3.39 (0.36)	−0.01	−4.09, −2.63
Exposure variables: X_1, X_2	−3.39 (0.35)	−0.01	−4.07, −2.69
G-Computation			
Super learner			
Outcome variables: A, X_1, X_2, X_3	−3.27 (0.35)	0.11	−3.98, −2.56
Misspecified parametric regression			
Main-terms misspecification			
Outcome variables: A, X_1, X_2, X_3	−3.25 (0.33)	0.13	−3.91, −2.59
Omitted-variable misspecification			
Outcome variables: A, X_1, X_2	−4.98 (0.37)	−1.60	−5.69, −4.24[b]
Inverse Probability Weighting			
Super learner			
Exposure variables: X_1, X_2, X_3	−3.43 (0.37)	−0.05	−4.17, −2.63
Misspecified parametric regression			
Omitted-variable misspecification			
Exposure variables: X_1, X_2	−4.96 (0.37)	−1.58	−5.67, −4.21[b]

Takeaways

TMLE Advantages

- Doubly robust
- Outperforms G-comp and IPW under model misspecification
- Substitution estimator
- Easily implemented in R
- Best practice: use ensemble machine learning

Overall: TMLE is a powerful and flexible estimator for causal inference in observational studies

American Journal of Epidemiology
© The Author(s) 2020. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.
Vol. 189, No. 8
DOI: 10.1093/aje/kwz277
Advance Access publication:
January 13, 2020

Original Contribution

US Migration Status of Adult Children and Cognitive Decline Among Older Parents Who Remain in Mexico

Jacqueline M. Torres*, Oleg Sofrygin, Kara E. Rudolph, Mary N. Haan, Rebeca Wong, and M. Maria Glymour

* Correspondence to Dr. Jacqueline M. Torres, Department of Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, 550 16th Street, San Francisco, CA 94143 (e-mail: Jacqueline.Torres@ucsf.edu).

Stanford University

Background & Study Design

- Cognitive decline is a growing concern in LMICs, and Mexico is among those experiencing rapid population aging
- Prior studies found mixed results on effects of adult child migration
- This study is the first to examine cognitive decline (not just levels) in relation to US migration of adult children
- **Main hypothesis:** having an adult child in the US is associated with faster cognitive decline in older Mexican parents

Study Design

- **Exposure:** Having an adult child in the US in both 2001 and 2003
- **Outcomes:**
  - Change in verbal memory z-scores (immediate + delayed recall)
  - Global cognitive performance
- **TMLE:** used to estimate marginal risk differences
  - Adjusts for time-varying confounding and attrition
  - Implemented with gradient boosting + Super Learner

TMLE in Practice

Why TMLE in this study?

- Migration is not random - confounded by social, economic, and health factors
- TMLE allows flexible modeling of exposure and outcome
- Handles time-varying confounding, attrition, and model misspecification

TMLE Approach

- **Purpose**: Estimate effect of adult child US migration on cognitive decline
  - Compared outcomes under two counterfactual scenarios
  - Adjusted for time-invariant and time-varying confounders
  - Handled loss to follow-up with attrition weights
  - Used machine learning (gradient boosting + Super Learner)

Table 2. Marginal Risk Differences in Average Change in Cognitive Performance z Scores for Adults Aged ≥50 Years With At Least 1 Adult Migrant Child at Baseline and 2-Year Follow-up Versus No Migrant Children, by Sex, Mexican Health and Aging Study, 2001–2015

Cognitive Outcome	Women[a]		Men[b]	
	Marginal RD	95% CI	Marginal RD	95% CI
9-year change				
Immediate verbal recall	−0.09	−0.156, −0.027	0.01	−0.062, 0.072
Delayed verbal recall	−0.10	−0.165, −0.031	−0.03	−0.106, 0.046
Global cognitive score	−0.04	−0.068, −0.004	−0.01	−0.051, 0.034
12-year change				
Immediate verbal recall	−0.13	−0.211, −0.055	−0.04	−0.102, 0.017
Delayed verbal recall	−0.12	−0.183, −0.047	0.02	−0.049, 0.082
Global cognitive score	−0.02	−0.050, 0.011	−0.01	−0.036, 0.023

Abbreviations: CI, confidence interval; RD, risk difference.

[a] $n = 3{,}416$ women included in estimates of 9-year change scores; $n = 2{,}838$ women included in estimates of 12-year change scores.

[b] $n = 2{,}556$ men included in estimates of 9-year change scores; $n = 2{,}101$ men included in estimates of 12-year change scores.

Takeaways

Discussion & TMLE Relevance

- Family member migration may be a social determinant of cognitive aging in LMICs
- TMLE enabled doubly robust, longitudinal estimation!
  - Adjusted for time-varying confounding and attrition
  - Used machine learning (gradient boosting and Super Learner)
- Findings were sex-specific, effects observed primarily among women
- Authors suggest future work should investigate mechanisms

Discussion Questions

Schuler and Rose

1. TMLE relies on correctly specifying either the outcome or treatment model. In practice, how do we assess which of the models is well specified?
2. The targeting step updates predictions using the updated or clever covariate. Conceptually, how should we interpret this update — as bias correction, efficiency gain, or something else?
3. In the simulation, TMLE outperforms others under misspecification. But how might its performance change with real-world challenges like weak positivity or limited overlap?

Torres

1. Are the causal assumptions reasonable in this migration context?
2. I was curious about this - why do people think the effect showed up only for women?

Citations

Megan S. Schuler, Sherri Rose, Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies, American Journal of Epidemiology, Volume 185, Issue 1, 1 January 2017, Pages 65–73, https://doi.org/10.1093/aje/kww165

Jacqueline M Torres, Oleg Sofrygin, Kara E Rudolph, Mary N Haan, Rebeca Wong, M Maria Glymour, US Migration Status of Adult Children and Cognitive Decline Among Older Parents Who Remain in Mexico, American Journal of Epidemiology, Volume 189, Issue 8, August 2020, Pages 761–769, https://doi.org/10.1093/aje/kwz277

Thank You!