Targeted learning to generate real world evidence

Mark van der Laan

Jiann-Ping Hsu/Karl E. Peace Professor in Biostatistics & Statistics
University of California, Berkeley

ISPOR, May 6, 2024, Atlanta

Acknowledgements: Maya Petersen, Rachael Phillips, Susan Gruber, Ivana Malenica

Poll Question 1

It's Time for a Poll!

1. What are your biggest concerns with RWE as a source of evidence for causal inference?

a) No concerns, it helps address gaps from clinical trials

b) Concerns with confounding

c) Concerns with acceptability by decision makers

d) Concerns with lack of fit-for-purpose data sources

e) Other concerns

Poll Question 2

It's Time for a Poll!

2. How familiar are you with causal inference methods for real world data?

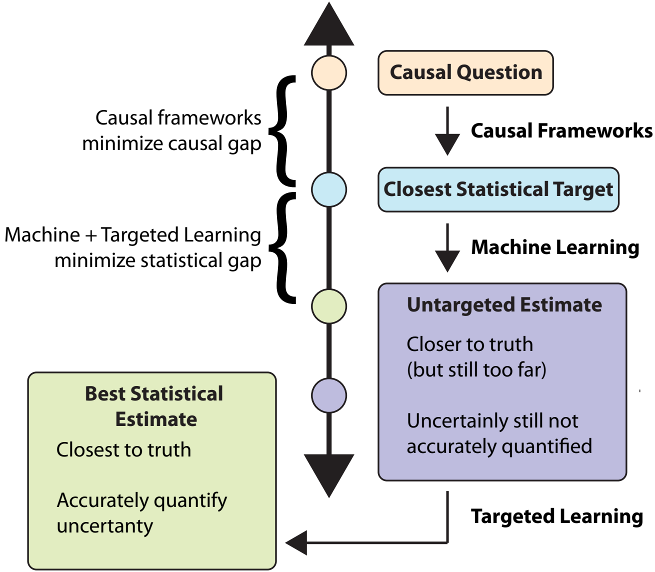
a) Not at all familiar

b) Somewhat familiar

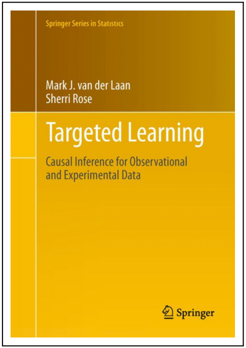
c) Very familiar with traditional methods

d) Very familiar with traditional and doubly-robust machines learning based methods such as augmented inverse probability of treatment weighting and targeted maximum likelihood estimation (A-IPTW/TMLE)

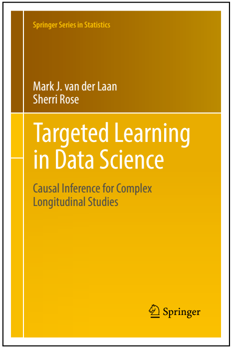
Targeted Learning for answering statistical and causal questions with confidence intervals



Targeted Learning is a subfield of statistics



van der Laan & Rose, *Targeted Learning: Causal Inference for Observational and Experimental Data*. New York: Springer, 2011.



van der Laan & Rose, *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*. New York: Springer, 2018.

The Hitchhiker's Guide to the `tlverse`

Better clinical decisions from observational data

Statistics in Medicine

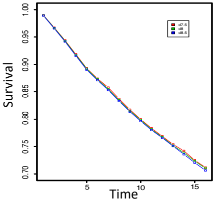
Research Article

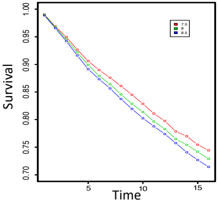
Received 24 May 2013, Accepted 5 January 2014 Published online 17 February 2014 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6099

Targeted learning in real-world comparative effectiveness research with time-varying interventions

Romain Neugebauer,[a]*† Julie A. Schmittdiel[a] and Mark J. van der Laan[b]

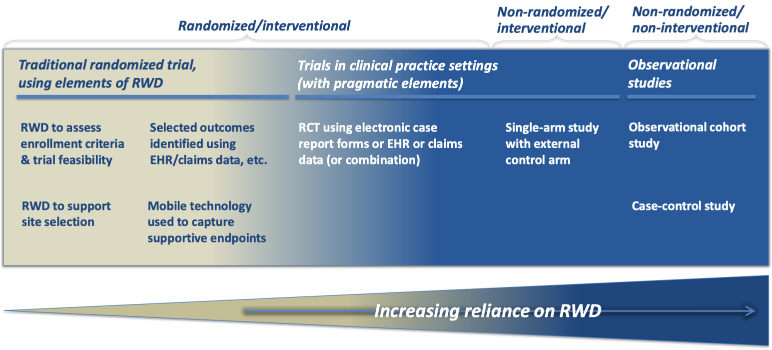




Standard methods: No benefit to more aggressive intensification strategy

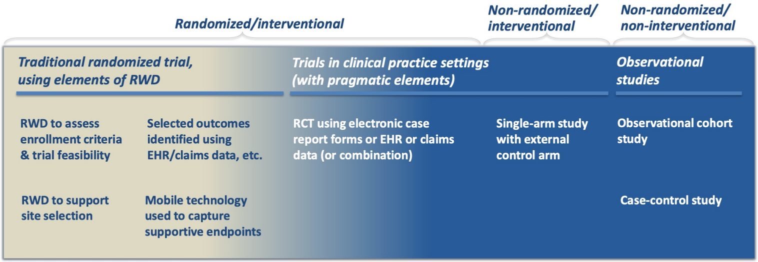
Targeted Learning: More aggressive intensification protocols result in better outcomes

Statistical challenges with RWD



Courtesy of "FDA Real-World Evidence Program" Webinar by John Concato on 4 August 2021

Statistical challenges with RWD



RWD Challenges

- ❑ Selection bias
- ❑ Intercurrent events
- ❑ Informative missingness
- ❑ Treatment by indication
- ❑ High dimensional covariates
- ❑ Outcome measurement error
- ❑ Statistical model misspecification
- ❑ Differences between external controls and single trial arm RCT

The roadmap for learning from data

STEP 1:
DESCRIBE EXPERIMENT

STEP 2:
SPECIFY STATISTICAL MODEL

STEP 3:
DEFINE STATISTICAL QUERY

STEP 4:
CONSTRUCT ESTIMATOR

STEP 5:
OBTAIN INFERENCE

STEP 6:
MAKE SUBSTANTIVE CONCLUSION

What is the experiment that generated the data?

STEP 1: DESCRIBE EXPERIMENT

STEP 2: SPECIFY STATISTICAL MODEL

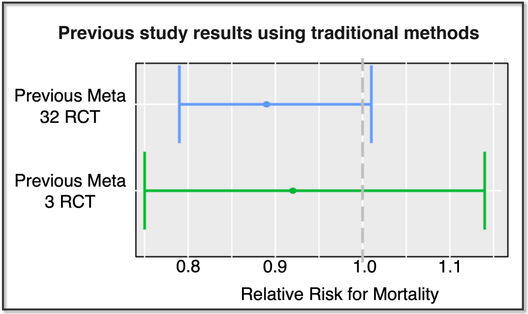
STEP 3: DEFINE STATISTICAL QUERY

STEP 4: CONSTRUCT ESTIMATOR

STEP 5: OBTAIN INFERENCE

STEP 6: MAKE SUBSTANTIVE CONCLUSION

Three multi-national RCTs assessing impact of corticosteroids on mortality among septic shock patients



What is the experiment that generated the data?

STEP 1: DESCRIBE EXPERIMENT

STEP 2: SPECIFY STATISTICAL MODEL

STEP 3: DEFINE STATISTICAL QUERY

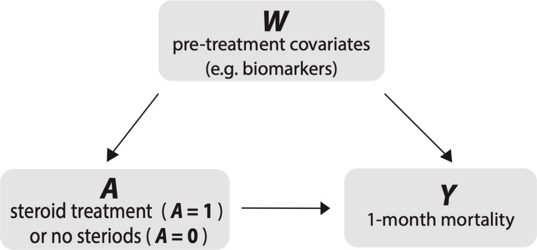
STEP 4: CONSTRUCT ESTIMATOR

STEP 5: OBTAIN INFERENCE

STEP 6: MAKE SUBSTANTIVE CONCLUSION

Three multi-national RCTs assessing impact of corticosteroids on mortality among septic shock patients

Pooled sample of $n = 1{,}300$ adults in septic shock



What is known about stochastic relations of the observed variables?

STEP 1:
DESCRIBE
EXPERIMENT

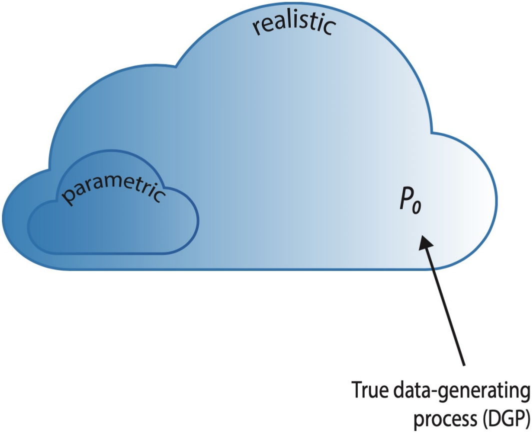
**STEP 2:
SPECIFY
STATISTICAL MODEL**

STEP 3:
DEFINE STATISTICAL
QUERY

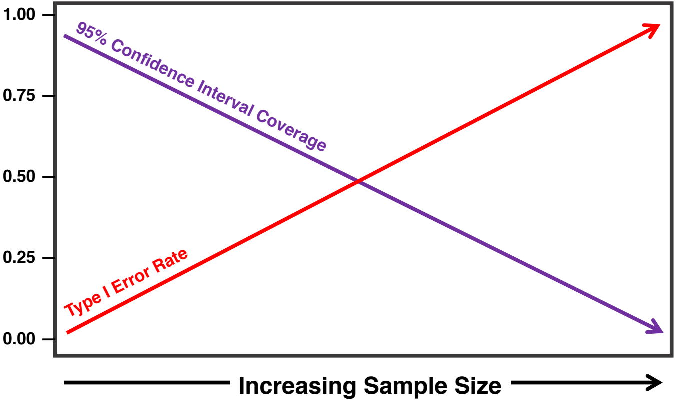
STEP 4:
CONSTRUCT
ESTIMATOR

STEP 5:
OBTAIN
INFERENCE

STEP 6:
MAKE SUBSTANTIVE
CONCLUSION



What happens when the statistical model is misspecified and does not contain the DGP?



Step 3a: What is the target causal estimand that we aim to identify from the data?

- STEP 1: DESCRIBE EXPERIMENT
- STEP 2: SPECIFY STATISTICAL MODEL
- **STEP 3: DEFINE STATISTICAL QUERY**
- STEP 4: CONSTRUCT ESTIMATOR
- STEP 5: OBTAIN INFERENCE
- STEP 6: MAKE SUBSTANTIVE CONCLUSION

What is the causal risk difference in mortality between treatment groups?

$$\psi_{causal} = E[Y_1 - Y_0]$$

Step 3b: What is the target statistical estimand that we will learn from the data?

STEP 1: DESCRIBE EXPERIMENT

STEP 2: SPECIFY STATISTICAL MODEL

STEP 3: DEFINE STATISTICAL QUERY

STEP 4: CONSTRUCT ESTIMATOR

STEP 5: OBTAIN INFERENCE

STEP 6: MAKE SUBSTANTIVE CONCLUSION

What is the average difference in mortality between treatment groups when adjusting for covariates?

$$\psi_{stat} = E(E[Y|A=1,W] - E[Y|A=0,W])$$

How should we estimate the target estimand?

STEP 1: DESCRIBE EXPERIMENT

STEP 2: SPECIFY STATISTICAL MODEL

STEP 3: DEFINE STATISTICAL QUERY

STEP 4: CONSTRUCT ESTIMATOR

STEP 5: OBTAIN INFERENCE

STEP 6: MAKE SUBSTANTIVE CONCLUSION

Statistical properties to consider

- Substitution / plug-in
- Valid inference
- Efficiency
- Ability to optimize finite sample performance

Targeted Maximum Likelihood Estimation (TMLE)

STEP 1: DESCRIBE EXPERIMENT

STEP 2: SPECIFY STATISTICAL MODEL

STEP 3: DEFINE STATISTICAL QUERY

STEP 4: CONSTRUCT ESTIMATOR

STEP 5: OBTAIN INFERENCE

STEP 6: MAKE SUBSTANTIVE CONCLUSION

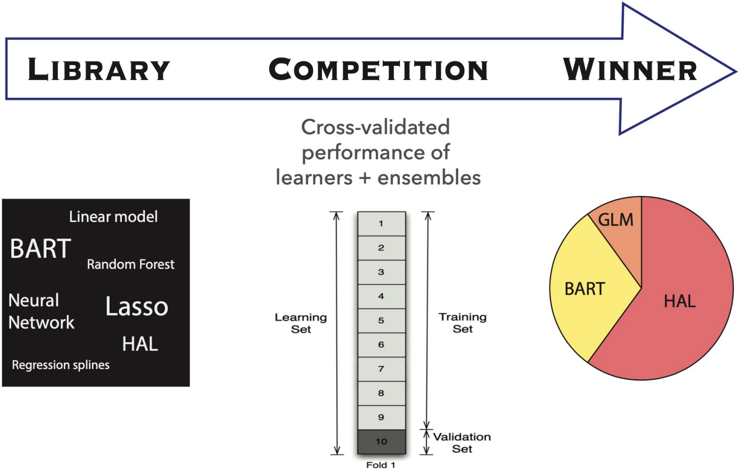
TMLE

1. Initial estimation of $E[Y|A,W]$ with super (machine) learning
2. Updating initial estimate to acheive optimal bias-variance trade-off for ψ_{stat}

TMLE estimates are optimal:

plug-in, efficient, unbiased, finite sample robust

TMLE Step 1: Super learner



Hugely advantageous when coupled with NLP-derived covariates with EHR

Highly Adaptive Lasso (HAL)

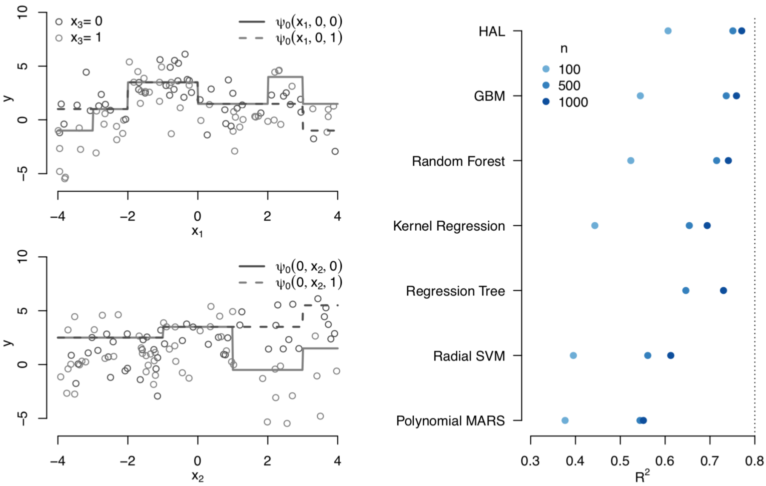
Key Idea

- Any d-dimensional cadlag function (i.e. right-continuous) can be represented as a possibly infinite linear combination of spline basis functions.
- The variation norm / complexity of a function is the L_1-norm of the vector of coefficients.

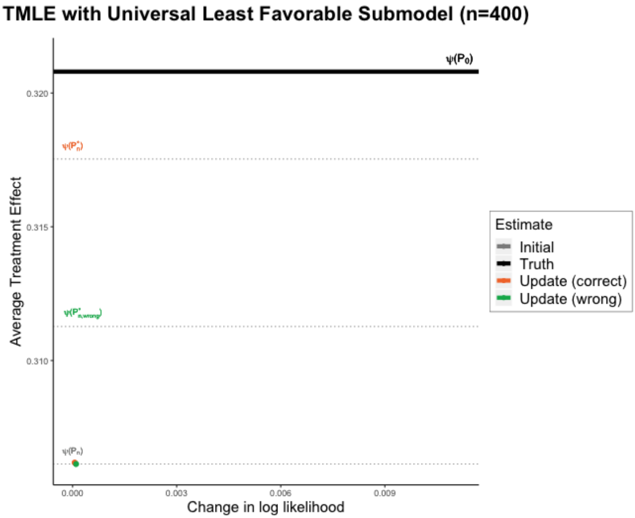
Converges to true function at rate $n^{-1/3}(\log n)^{d/2}$

HAL performance for d=3

$$\psi_0(x) = -2x_3\, I(x_1 < -3) + 2.5\, I(x_1 > -2) - 2\, I(x_1 > 0) + 2.5x_3\, I(x_1 > 2) -2.5\, I(x_1 > 3) + I(x_2 > -1) - 4x_3\, I(x_2 > 1) + 2\, I(x_2 > 3)$$



TMLE Step 2: Targeting follows a path of maximal change in target estimand per unit likelihood



How should we approximate the sampling distribution of our estimator?

STEP 1: DESCRIBE EXPERIMENT

STEP 2: SPECIFY STATISTICAL MODEL

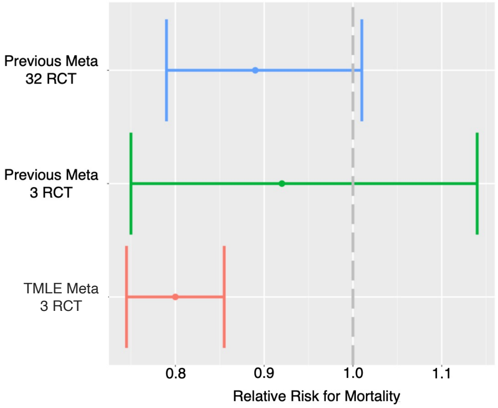
STEP 3: DEFINE STATISTICAL QUERY

STEP 4: CONSTRUCT ESTIMATOR

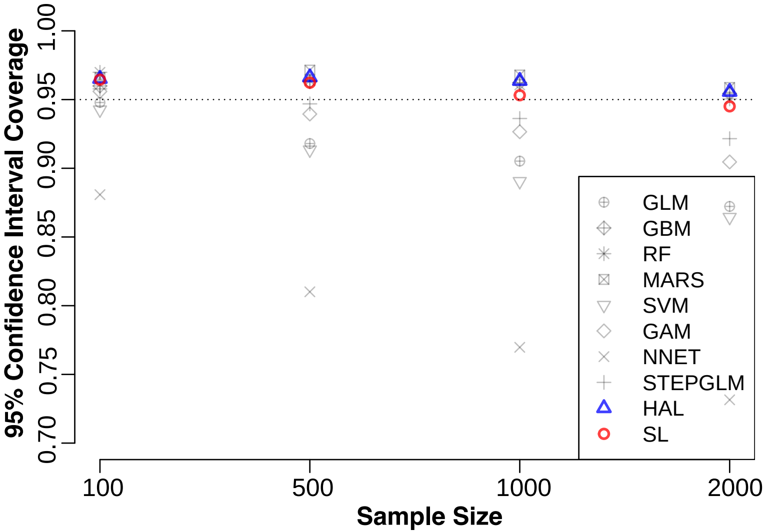
STEP 5: OBTAIN INFERENCE

STEP 6: MAKE SUBSTANTIVE CONCLUSION

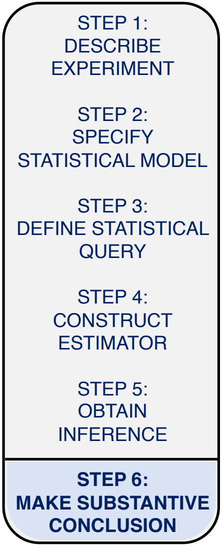
Due to targeting (step 2), the TMLE behaves as the ***sample mean*** of efficient influence function



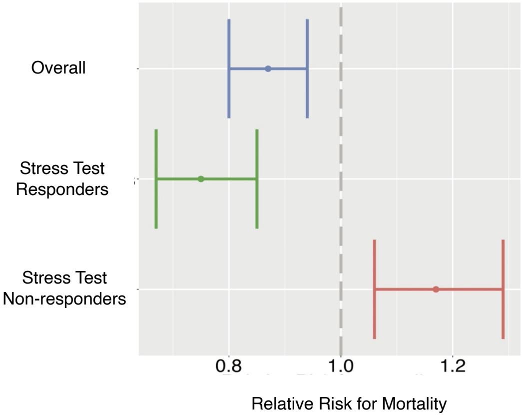
Can we break HAL-TMLE?



Possibility to refine question of interest and inform future studies



What subgroup of patients in septic shock benefit from corticosteroids?



Arriving at the substantive conclusion

STEP 1: DESCRIBE EXPERIMENT

STEP 2: SPECIFY STATISTICAL MODEL

STEP 3: DEFINE STATISTICAL QUERY

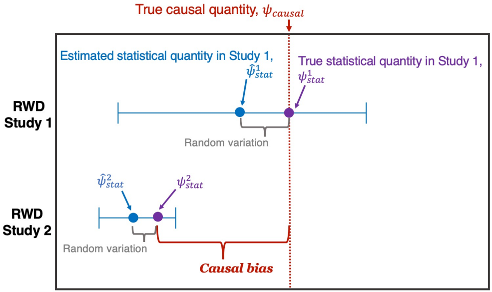
STEP 4: CONSTRUCT ESTIMATOR

STEP 5: OBTAIN INFERENCE

STEP 6: MAKE SUBSTANTIVE CONCLUSION

Investigate causal bias with sensitivity analysis

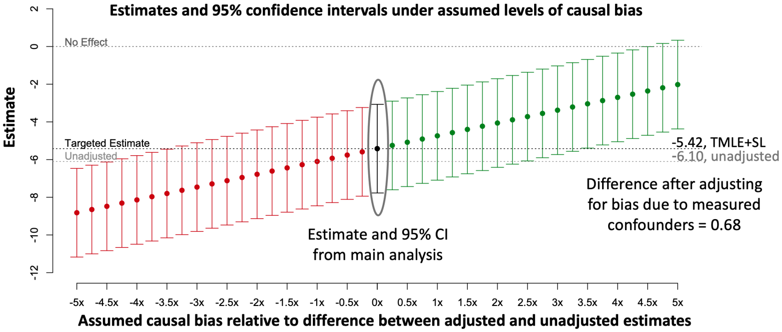
Causal bias: Gap between estimate and truth due to violations of any of the causal assumptions (e.g., unmeasured confounding)*



Sensitivity Analysis: Model-free assessment of how reasonable departures from causal assumptions would impact study findings

* Sensitivity analysis can be extended to incorporate statistical bias

TL-based non-parametric sensitivity analysis RCT with 25% LTFU example



Courtesy of "Targeted-Learning Based Statistical Analysis Plan" Webinar by Susan Gruber on 28 April 2021

Targeted Learning with RWD

Randomized/interventional			*Non-randomized/ interventional*	*Non-randomized/ non-interventional*
Traditional randomized trial, using elements of RWD		***Trials in clinical practice settings (with pragmatic elements)***		***Observational studies***
RWD to assess enrollment criteria & trial feasibility	Selected outcomes identified using EHR/claims data, etc.	RCT using electronic case report forms or EHR or claims data (or combination)	Single-arm study with external control arm	Observational cohort study
RWD to support site selection	Mobile technology used to capture supportive endpoints			Case-control study

RWD Challenges

- ❑ Selection bias
- ❑ Intercurrent events
- ❑ Informative missingness
- ❑ Treatment by indication
- ❑ High dimensional covariates
- ❑ Outcome measurement error
- ❑ Statistical model misspecification
- ❑ Differences between external controls and single trial arm RCT

Targeted Learning path supports regulatory decision making

Targeted Learning

- ✓ Roadmap for causal and statistical inference
- ✓ Realistic statistical model
- ✓ Statistical estimand approximates answer to causal question
- ✓ Flexible estimation and dimension reduction with Super Learner
- ✓ Model-free sensitivity analysis
- ✓ Generate RWE with confidence

Concluding Remarks

- Roadmap for causal inference and Targeted Learning provides systematic principled approach for generating RWE.
- Integrates all advances in machine learning, statistical theory and causal identification.
- SL and TMLE can be tailored towards particular estimation problem in pre-specified manner using outcome blind simulations.

Poll Question 3

It's Time for a Poll!

3. What is the key feature of doubly robust methods based on machine learning?

a) Only one model fits (among treatment/censoring mechanism, and outcome regression) needs to be correct

b) Super learning ends up with better algorithm for fitting these regressions than any one algorithm

c) Doubly-robust ML-based methods have a higher likelihood of getting a correct effect estimate

d) All of the above