Trials Designs for Adaptive Interventions –Research Questions Closer to Practice in Trials

Maya Petersen
Div. Epidemiology & Biostatistics
School of Public Health,
University of California, Berkeley

Precision Public Health

- Precision Medicine (NIH):
  - "An emerging approach for disease treatment and prevention that **takes into account individual variability** in genes, environment, and lifestyle for each person."
- Concept also central to **optimizing the impact of public health interventions**
  - Improve outcomes for more people
  - Improve outcomes for as many people as possible given limited resources

https://www.nih.gov/precision-medicine-initiative-cohort-program

Precision Public Health

1. Improve outcomes for more people
   - **Variability in effectiveness** of interventions
     - Variability across individuals, clinics, communities, contexts,…
   - Give each person the intervention he/she most likely to benefit from
2. Improve outcomes for as many people as possible given limited resources
   - **Variability in underlying risk of a poor outcome**
   - Reserve costly interventions for those who both need them and are likely to respond

https://www.nih.gov/precision-medicine-initiative-cohort-program

Precision Public Health

1. Improve outcomes for more people
   - **Variability in effectiveness** of interventions
     - Variability across individuals, clinics, communities, contexts,…
   - Give each person to the intervention he/she most likely to benefit from

“Adaptive Interventions”

AKA: Individualized treatments or “dynamic regimes”

outcome

– Reserve costly interventions for those who both need them and are likely to respond

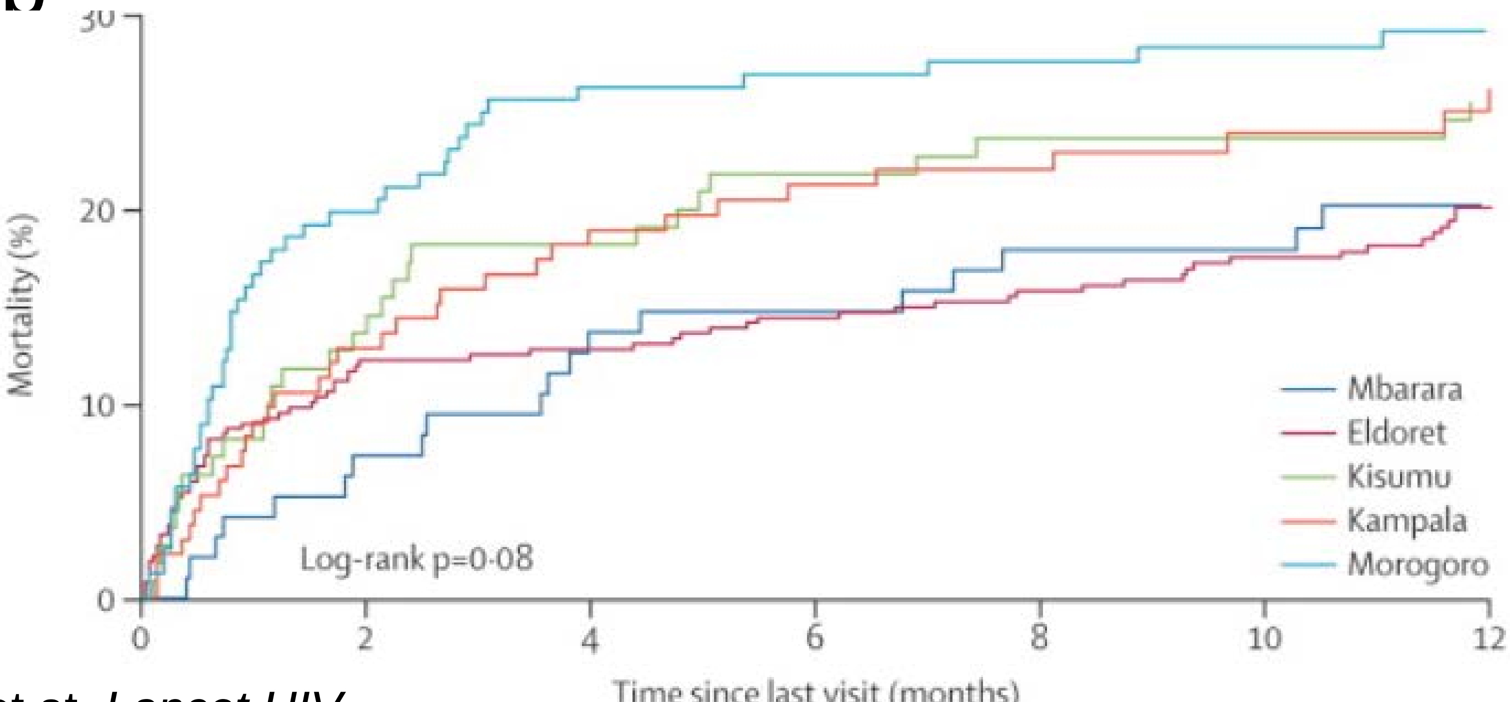
But don’t we use adaptive interventions all the time in practice?

- Yes!
- **But** we DON’T typically design or analyze studies with this goal in mind. And we should!
- Novel designs and novel analytic methods directly targeted at

1. Developing adaptive interventions that will give the best overall outcomes
2. Evaluating the comparative effectiveness of these adaptive interventions

Ex. Retention in HIV Care in East Africa

- Loss to follow up after enrollment in HIV care: 20-40% by two years
- High mortality among those lost to follow up



Geng et at, *Lancet HIV*,

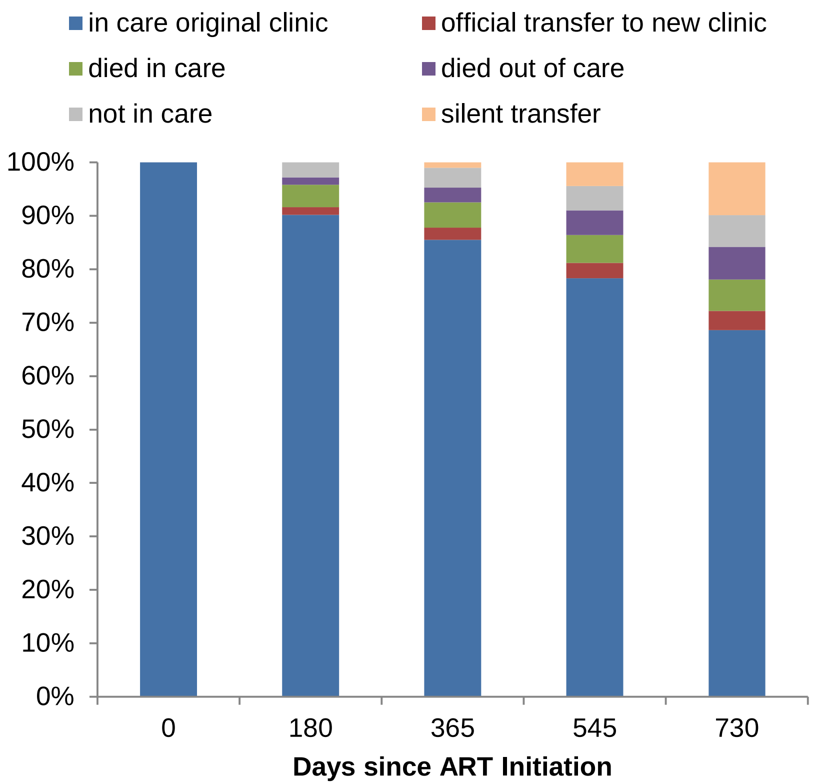
Interventions to improve retention in HIV Care

- Strategies to optimize retention within resource constraints urgently needed
- Several interventions with randomized trials showing effica

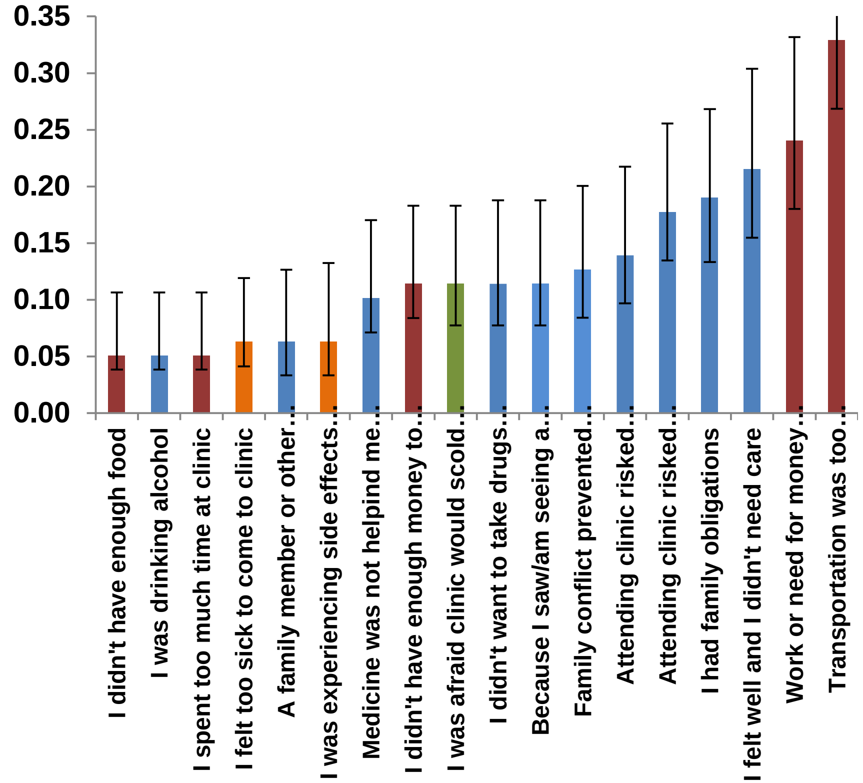
  - SMS Text message
    - Appointment reminders and build relationship
  - Transport vouchers
    - Small cash incentives for on time clinic visit
  - Peer Navigators
    - Peer health workers to navigate barriers

Need differs across patients and over time

Most patients stay in care with no intervention



Reasons for dropout vary



Geng et al, *CID*, 2016

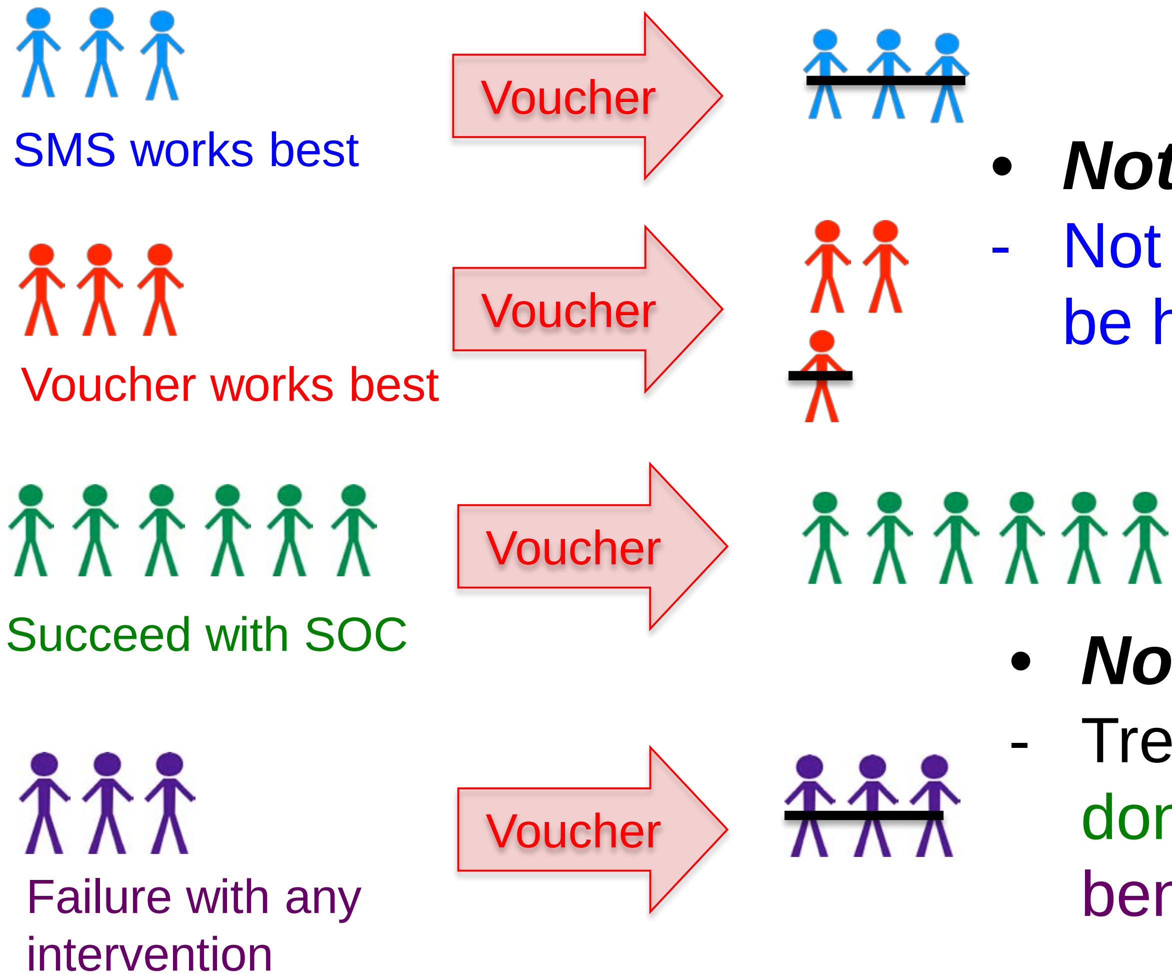
Traditional RCT Paradigm

- Active arm(s) versus standard of care (SOC)
- Example
  - Design: Randomize patients to eg. vouchers vs. SMS vs. standard-of-care (SOC)
  - Question: How would proportion retained (eg 2 years later) differ if everyone got a voucher vs. everyone got SOC?
    - "Average treatment effect"- compares "static" interventions
  - Analysis: Compare mean outcomes between arms
    - +/- some adjustment for precision
- Limitation: Average population effects may hide key heterogeneity in response

Limitations of "static" interventions

- "Static": All patients get the same intervention

Retention success:



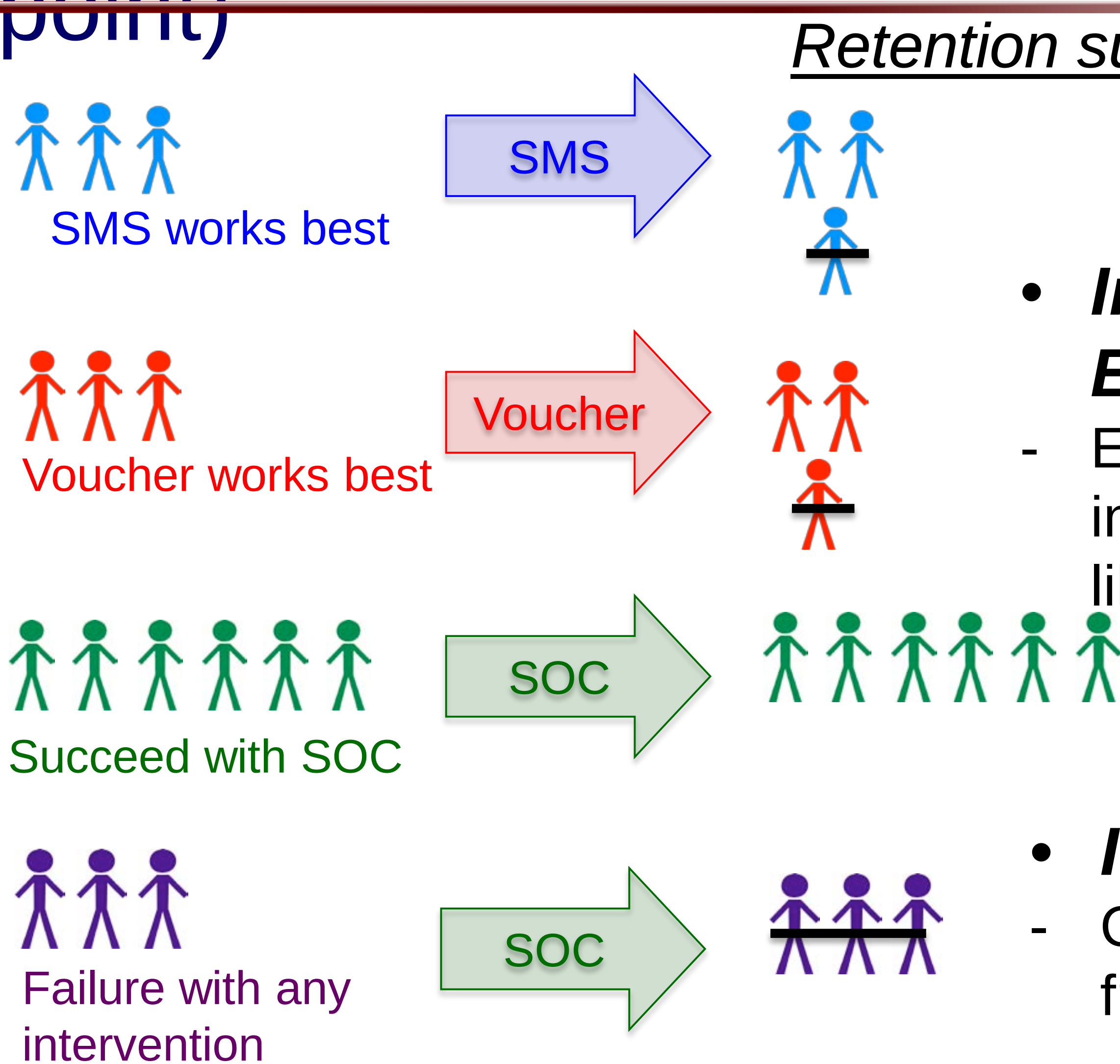
- ***Not optimally effective***
  - Not helping all who could be helped
- ***Not optimally efficient***
  - Treating patients who don't need or won't benefit from intervention

Beyond static interventions…

- How to better allocate our existing toolkit of interventions?
  - What is most effective/cost effective way to "tailor": i.e. assign and modify interventions based on evolving patient characteristics?
- **Adaptive intervention:** Rule for assigning and modifying an intervention based on individual (or clinic, community, …) observed past
  - Baseline and/or time varying characteristics

Review of DTR literature and methods: Dynamic Treatment Regimes in Practice: Planning Trials and Analyzing Data for Personalized Medicine, Moodie E and Kosorok M, eds, 2016.

Adaptive interventions can improve effectiveness and efficiency (single time point)



- ***Improved Effectiveness***
  - Each patient gets the intervention he/she most likely to benefit from

- ***Improved Efficiency***
  - Only those who will benefit from an intervention get it

SOC= "Standard of Care"

"Wait a second…Isn't this just a fancy way of discussing subgroup analyses of RCTs?"

- Traditional RCT approach to heterogeneity:
  - Pick a few *a priori* subgroups (not too many!)
  - Estimate average treatment effect for each
- Ex. Average effect of vouchers vs. SOC on retention among those who live far from vs. near to clinic…
  - And perhaps only see effect among those living far…
- <u>Limitations</u>
  - Which subgroups to choose? Might not know *a priori*
    - How to define "far"? Does living "far" only matter if also

Machine Learning to develop and evaluate optimal adaptive intervention strategies

- Which rule for assigning interventions would result in the highest retention?
  - – Super Learning
    - Learn optimal rule for assigning an initial intervention based on measured characteristics at baseline
    - Specific loss function- targeted at optimizing outcome
- What would outcomes have been if all patients had followed this rule
    - (vs. for example, all gotten vouchers or SMS)?
  - – Cross-validated Targeted Maximum Likelihood

Luedtke & van der Laan 2014; van der Laan and Luedtke 2014;

Nice in theory, but….

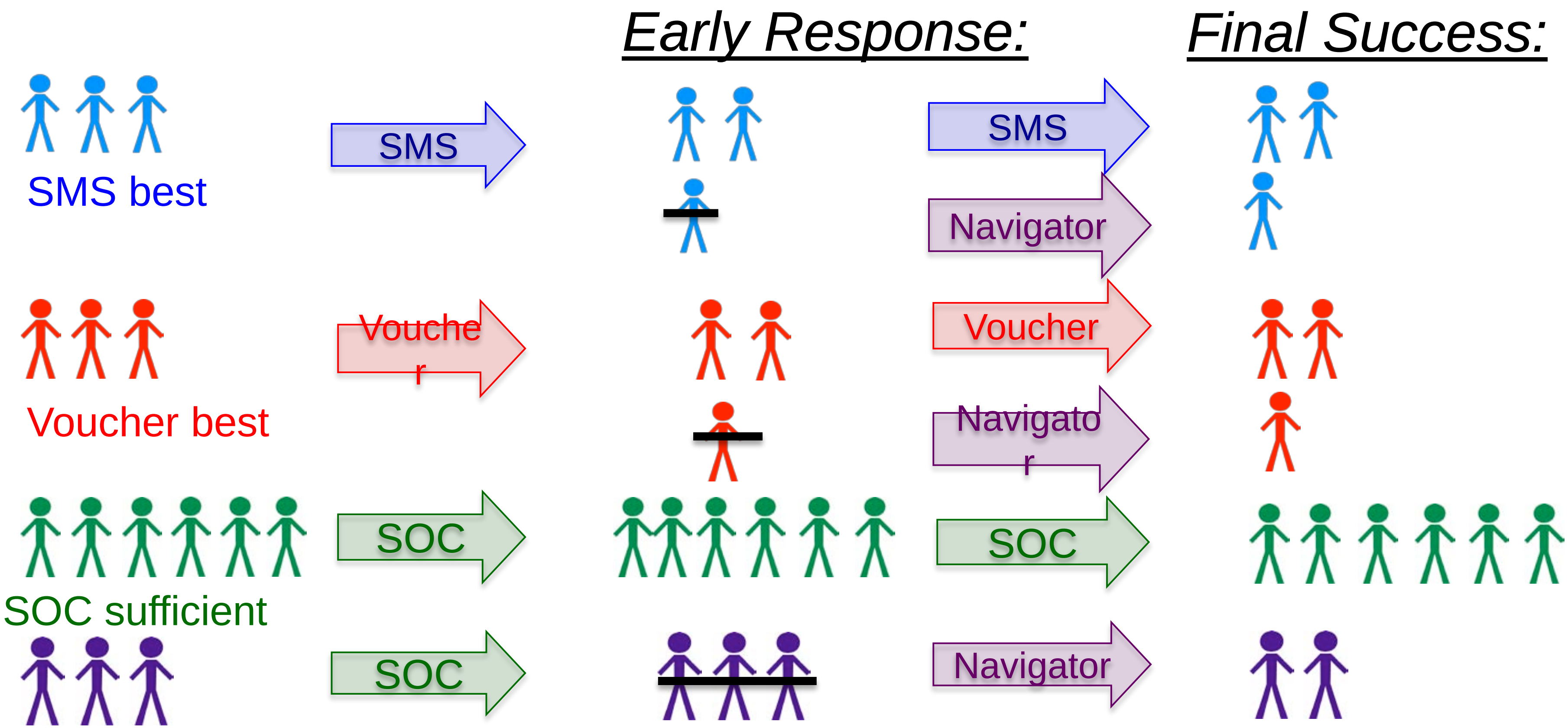
- Requires measuring patient characteristics that accurately predict response
- Ex: Can we actually measure enough on people to distinguish those who require vouchers from those who will do fine with SOC?
  - Maybe, maybe not….
- Using a patient’s own response to an initial intervention can help …

Longitudinal adaptive interventions offer additional advantages

1. Low cost/low intensity intervention at baseline
   - With or without additional targeting using baseline characteristics
2. Escalate to higher cost/intensity for those with early poor response
   - With or without additional targeting using time updated characteristics

- Advantages
  - **Effectiveness**: “salvage” when low intensity intervention insufficient, needs change, or imperfectly targeted
  - **Efficiency**: higher intensity intervention reserved for those with demonstrated need

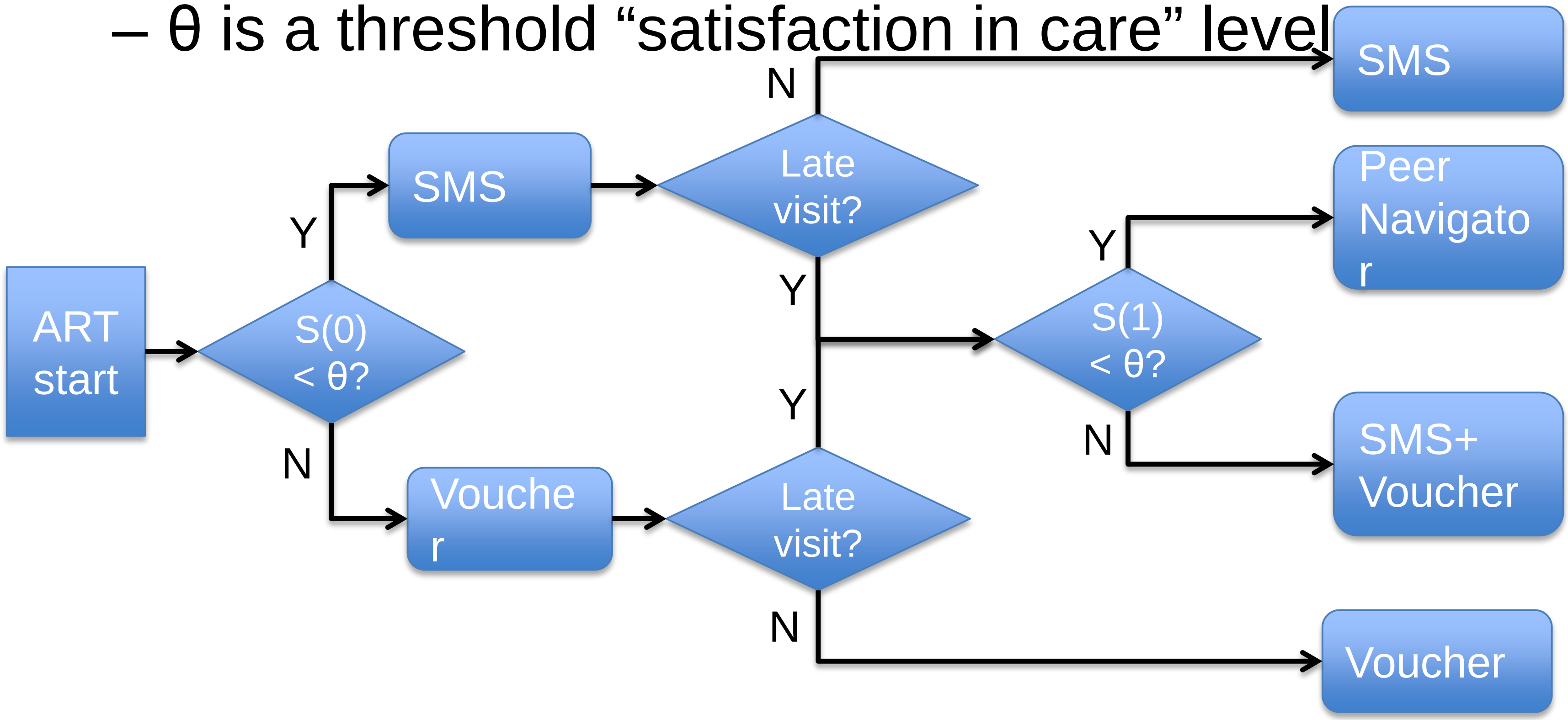
Ex: Longitudinal adaptive interventions



- ***2^{nd} line "salvage" further Improves effectiveness***
  - Patients who don't respond to early low cost intervention still helped
- ***Efficiency***

Ex: Using patient characteristics to assign treatment/modify interventions over time

- Rule d_θ for assigning and modifying interventions
  - Satisfaction with care
    - Marker for structural vs. psychosocial barriers to retention
    - Measured at ART start ($S(0)$) and 1st late visit ($S(1)$)
  - θ is a threshold "satisfaction in care" level



Goals (target causal parameters) for precision public health

1. **Expected outcome under a specific adaptive intervention**
   - Mean outcome if all subjects had followed a given rule for assigning and modifying interventions?

- Retention example
  - Outcome Y: Indicator retention 2 years after starting ART
  - Counterfactual outcome under rule d_θ : $Y(\theta)$
  - Goal: Estimate $E(Y(\theta))$ for some θ
    - Proportion of patients retained if all had followed rule d_θ
  - Effect relative to SOC: $E(Y(\theta)-Y(SOC))$

Goals (target causal parameters) for precision public health

2. Optimal adaptive intervention

– What rule would result in best mean outcome if all subjects followed it?

- Retention Example:
  – Rule d^{opt} for assigning intervention that would maximize proportion retained
    - Or maximize the proportion retained under resource constraints

1. Among all possible adaptive interventions?
   – i.e. Rules with access to all measured variables (or a subset deemed reasonably accessible in practice)
2. Among a specified subset of rules?
   – Ex: optimal satisfaction threshold θ^{opt}?

Goals (target causal parameters) for precision public health

3. **Comparative effectiveness of optimal adaptive intervention**
   - Mean outcome if all subjects followed optimal rule: $E[Y(d^{opt})]$

- Retention Example:
  - Effect compared to standard of care:
    - $E[Y(d^{opt})- Y(SOC)]$
  - Effect compared to a simpler adaptive option:
    - Ex: d^*:
      - 1st line Voucher for all
      - 2nd line Navigator for all early failures
    - $E[Y(d^{opt})- Y(d^*)]$

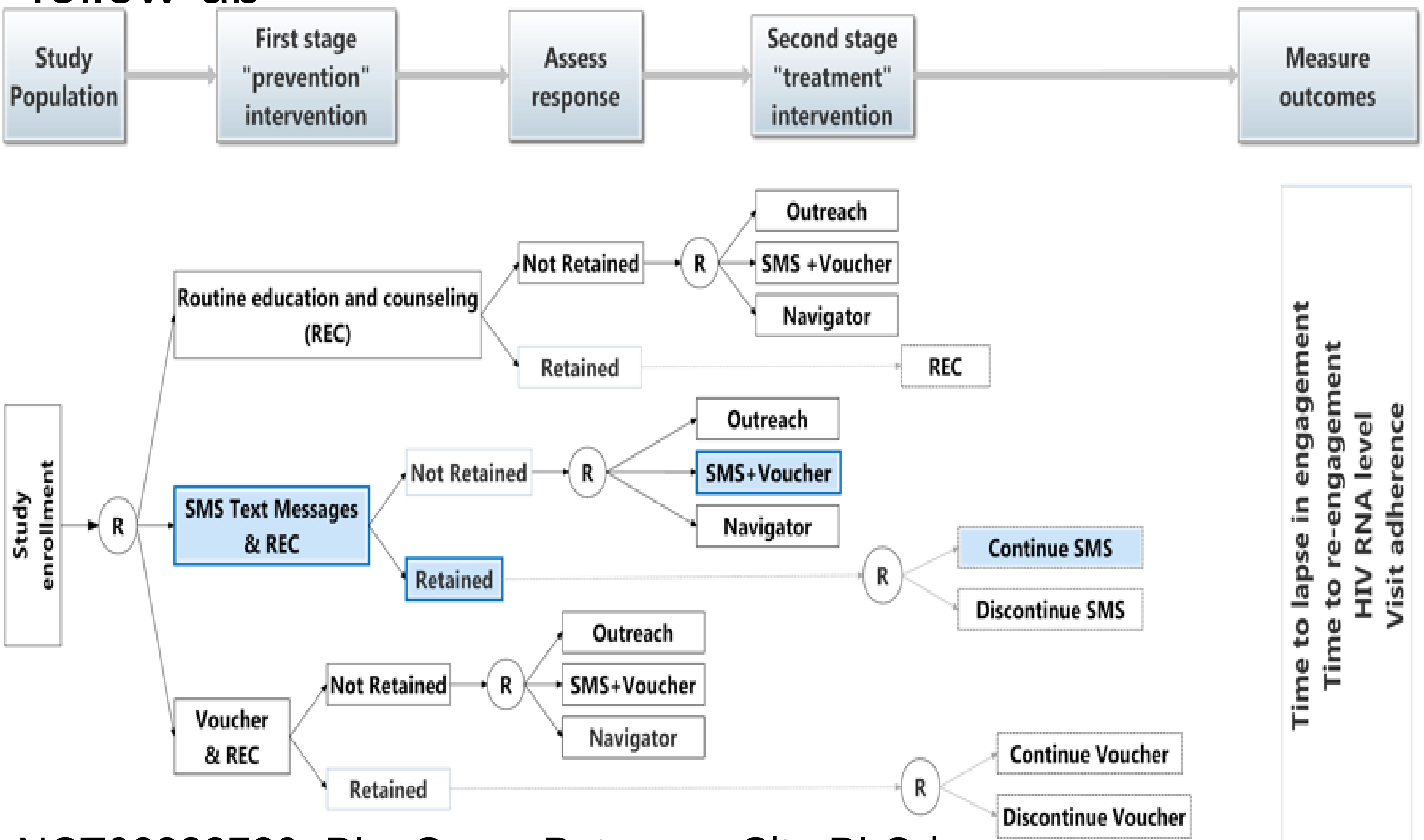
Experimental designs for building and evaluating longitudinal adaptive interventions

- "Sequentially Randomized Trials" or "Sequential Multiple Assignment Randomized Trials" (SMART designs)
- Define
  - Decision points for modifying an intervention
  - At baseline and each subsequent decision point, intervention options
    - Can depend on individual's observed past up to that time point
- At baseline and each time a decision is triggered, re-randomize intervention

See Murphy et al: Many references: https://methodology.psu.edu/ra/adap-inter

AdaPT-R: A sequential multiple assignment RCT

- ~1800 adult HIV patients starting ART in Kenya, 2 years follow-up



NCT02338739; PIs: Geng, Petersen; Site PI Odeny

Analytic methods for building and evaluating optimal longitudinal adaptive interventions

- Analytic methods: extensions from single time point methods
  - Super Learning
    - Learn optimal rule for each time point, sequentially from last time point (assuming future assignment follows optimal)
  - TMLE
    - Evaluate comparative effectiveness of the rule, with inference (95% CI and p values)
- Same methods can be applied to observational data
  - Assume no unmeasured confounding
    - Guaranteed by design in a SMART

Conclusions

- Opportunities
  - Big Data: big samples, lots of measures (including longitudinal data), diverse data types
    - Ex: Real time electronic adherence monitoring
    - Ex: Social network data
  - Targeted Machine Learning:

    Disparate high dimensional data -> optimal targeting strategies (dynamic regimes)
- Toward a precision public health paradigm
  - "Right intervention to the right patient/clinic/community at the right time"

COMING UP NEXT

- Smarter (faster, more efficient) designs to get us there
- Adaptive interventions ≠ Adaptive designs

Estimation: Expected outcome under a specific dynamic regime: E(Y(d))

1. Inverse probability weighting: known weights
   - A subject who follows rule gets weight:

     1/probability of following rule

     - Probability of following rule d if fail:

       1/3*1/3=1/9

     - Probability of following rule d if succeed on SMS or Voucher:

       1/3*1/2=1/6

     - Probability of following rule d if succeed on SOC:

       1/3*1=1/3

   - A subject who does not follow rule gets weight: 0
   - Take average of weighted outcome

Estimation: Expected outcome under a specific dynamic regime: E(Y(d))

- Because interventions randomized, additional adjustment not needed to control for confounding
  - Adjusting for additional predictors of outcome can reduce variance
- Here discuss two approaches to further adjustment
  - No risk of bias in SMART

1. Inverse probability weighting - estimated weights
2. Targeted Maximum Likelihood

IPW: Robins & Rotnitzky, 1992; Hernan et al., 2006;
TMLE: Bang & Robins, 2005; van der Laan & Gruber 2012

Estimation: Expected outcome under a specific dynamic regime: $E(Y(d))$

2. Inverse probability weighting: using estimated weights
   - Estimate treatment mechanism: probability of following rule at each time point given data measured up to that time point

3. Targeted Maximum Likelihood
   - Estimate treatment mechanism (weights)
   - Estimate series of iterated outcome regressions
   - Further efficiency gains

Estimation: Optimal dynamic regime

1. Evaluate directly:
   – Estimate E(Y(d)) for each candidate d
   – Choose the d the minimizes failure probability

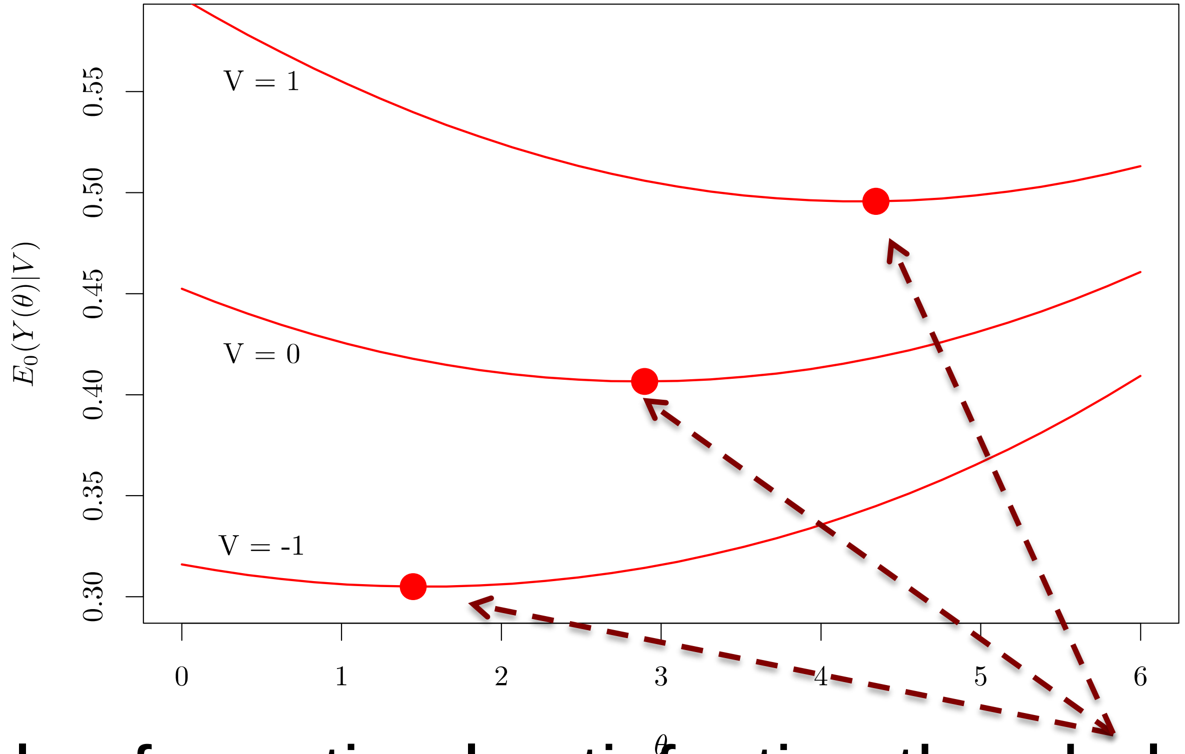
2. With a dynamic marginal structural model
   – Lower dimensional summary of how E(Y(d)) varies as a function of d
     • Possibly conditional on baseline covariates V
   – Ex. Model for how probability of failure depends on satisfaction threshold θ and baseline wealth V

MSM: Robins, 1999; Dynamic MSM: Petersen & van der Laan, 2007

Example: Dynamic Marginal Structural Model

- Model probability of failure given satisfaction threshold θ and baseline wealth V

  $E(Y(\theta)|V)=\text{expit}(\beta_0 + \beta_1\theta + \beta_2\theta^2 + \beta_3 V + \beta_4\theta V)$



- Solve for optimal satisfaction threshold θ given baseline wealth V (ie value that minimizes E(Y(θ)|V)):

  $\theta^{opt}(V)=\beta_1/2\beta_2 - \beta_4/2\beta_2 V$

Dynamic Marginal Structural Model

- Estimate of parameters β of marginal structural model yields estimate of
  1. Expectation under rule d^{θ} for some threshold θ (given V): $E(Y(\theta)|V)$
  2. Optimal Regime (within class):

     $\theta^{opt}(V)=\beta_1/2\beta_2-\beta_4/2\beta_2 V$
  3. Expected outcome if everyone followed optimal rule: $E(Y(\theta^{opt}(V)))$
     - Just estimate $E(Y(\theta))$, plugging in estimate of $\theta^{opt}(V)$

Zhang et al., 2013

Estimation: Dynamic Marginal Structural Model

- Estimators of β in marginal structural model: Analogous to estimators of E(Y(d))

1. Inverse probability weighted

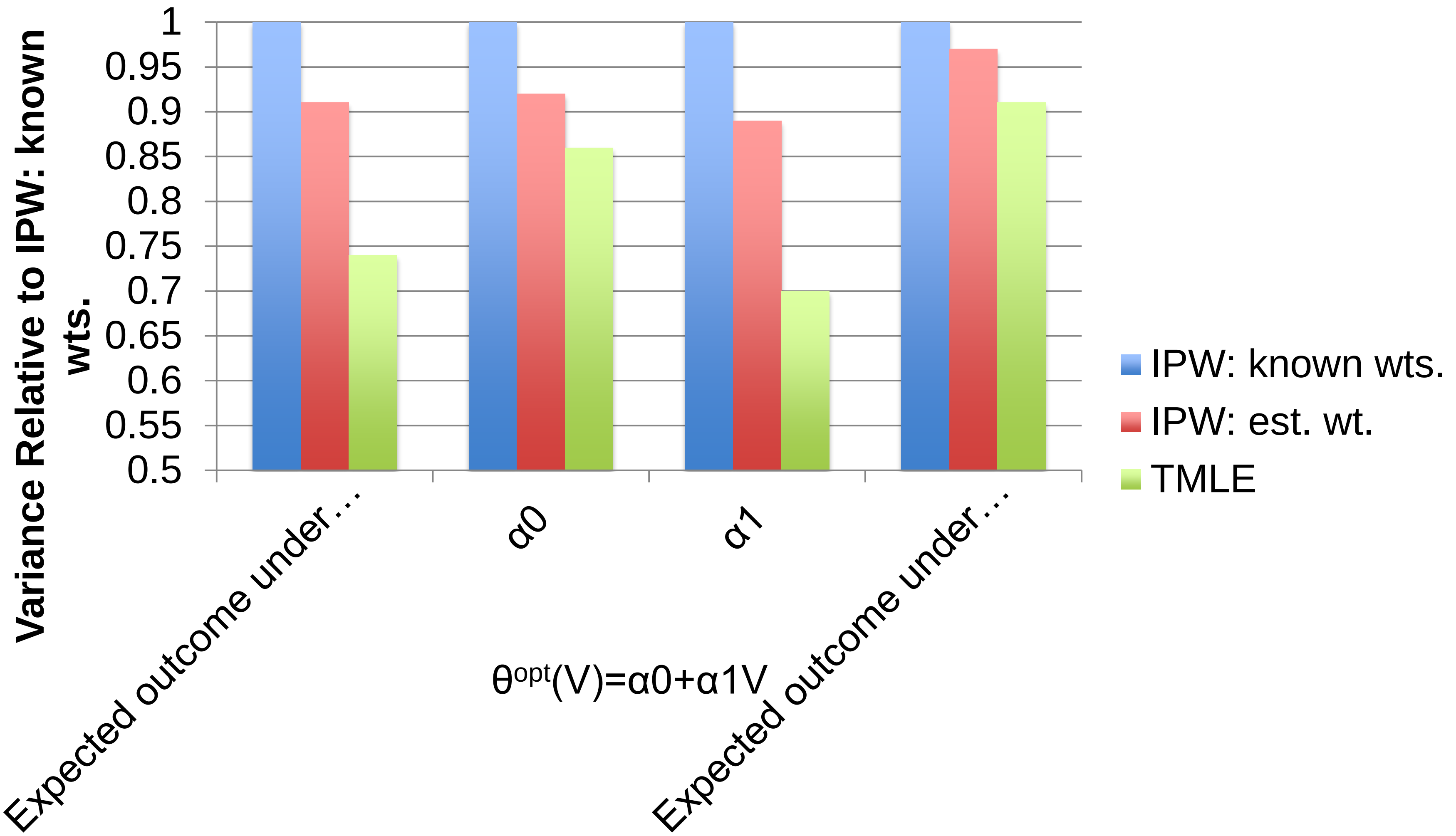
- Fit weighted regression with
  - Known weights – unbiased
  - Estimated weights- more efficient

2. Targeted Maximum Likelihood
   - Improve efficiency further

Robins, 1999; Petersen & van der Laan, 2007; Schnitzer et. al., 2013; Petersen et. al., 2014

Covariate adjustment reduces variance

All Estimators: good/conservative 95% CI coverage and Type I error control



Petersen et al. Ch 10 In: Moodie & Kosorok, Dynamic Treatment Regimes in Practice, 2016

Code & Simulated Data

- Code implementing examples here using ltmle R package:
  - Petersen et. al. Ch 10. In: Dynamic Treatment Regimes in Practice, Moodie E and Kosorok M, editors 2016
- ltmle R package
  - Causal effect estimation with multiple intervention nodes
    - Longitudinal static and dynamic regimes
    - Static and dynamic marginal structural working models
  - General longitudinal data structures
    - Repeated measures outcomes
    - Right censoring
  - Estimators
    - IPTW
    - ICE G-comp
    - TMLE
    - Options include nuisance parameter estimation via glm regression formulas or calling SuperLearner()
- Other DR software also available (tmle, …)

http://cran.r-project.org/web/packages/ltmle/; Schwab et al 2013

Selected References

1. H. Bang and J.M. Robins. Doubly-robust estimation in missing data and causal inference models. Biometrics, 61:962- 972, 2005.
2. M A Hernan, E Lanoy, D Costagliola, and J M Robins. Comparison of dynamic treatment regimes via inverse probability weighting. Basic & Clinical Pharmacology & Toxicology, 98:237242, 2006.
3. M Petersen, J. Schwab, E Geng, and M van der Laan. Evaluation of longitudinal dynamic regimes with and without marginal structural working models. In Moodie E and Kosorok M, editors, Dynamic Treatment Regimes in Practice: Planning Trials and Analyzing Data for Personalized Medicine., chapter 10, pp. 157-186. ASA-SIAM, 2016.
4. M.L. Petersen, J. Schwab, S. Gruber, N. Blaser, M. Schomaker, and M. van der Laan. Targeted maximum likelihood estimation for dynamic and static marginal structural working models. Journal of Causal Inference, 2(2), 2014.
5. J. Robins and A. Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. In AIDS Epidemiology, pp: 297-331. Springer, 1992.
6. J.M. Robins. Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference, volume 116 of IMA, pages 95-134. Springer, New York, NY, 1999.
7. J.M. Robins. Robust estimation in sequentially ignorable missing data and causal inference models. In Proceedings of the American Statistical Association on Bayesian Statistical Science, 1999, pages 6-10, 2000.
8. M.E. Schnitzer, Erica E.M. Moodie, and Robert W. Platt. Targeted maximum likelihood estimation for marginal time-dependent treatment effects under density misspecication. Biostatistics, 14(1):1{14, 2013.
9. M J Van der Laan and M L Petersen. Causal eect models for realistic individualized treatment and intention to treat rules. The International Journal of Biostatistics, 3, 2007.
10. M.J. van der Laan and S. Gruber. Targeted minimum loss based estimation of causal eects of multiple time point interventions. The International Journal of Biostatistics, 8(1):Article 8, 2012.
11. Baqun Zhang, Anastasios A. Tsiatis, Eric B. Laber, and Marie Davidian. Robust estimation of