Causal Effects based on Randomized Clinical and Intervention Trials

Alan Hubbard and Nicholas P. Jewell
Division of Biostatistics
School of Public Health
University of California, Berkeley

Workshop on the Design and Analysis of Clinical Trials

Oct 28, 2011

Collaborators

Mark van der Laan

Farid Jamshidian

Supported by NIAID

References

- M.J. van der Laan and S. Rose. *Targeted Learning*. Springer. 2011.
- M. Rosenblum, N. P. Jewell, M. J. van der Laan, S. Shiboski, A. van der Straten and N. Padian, "Analyzing direct effects in randomized trials with secondary intervention: An application to HIV prevention trials." *Journal of the Royal Statistical Association, Series A*, 172, 2009, 443-465.
- S. Shiboski, M. Rosenblum and N. P. Jewell, "The impact of secondary condom interventions on the interpretation of results from HIV prevention trials," *Statistical Communications in Infectious Diseases,* Vol. 2 : Iss. 1, Article 2.
- F. Jamshidian, A. Hubbard, N. P. Jewell, "Accounting for perception, placebo, ad unmasking effects in estimating treatment effects in randomized clinical trials," to appear in *Statistical Methods in Medical Research*.
- M. Backonja, A. Beydoun, K. R. Edwards et al., "Gabapentin for the symptomatic treatment of painful neuropathy in patients with diabetes mellitus. A randomized controlled trial," *JAMA*, 280, 1998, 1831-1836.
- J. Simes, M. Voysey, R.O'Connell et al., "A novel method to adjust efficacy estimates for uptake of other active treatments in long-term clinical trials," PLoS ONE 5(1): e8580. doi:10.1371/journal.pone.0008580

Benefits of Applying Insights of Causal Inference to Clinical Trials

- The application of "causal inference" techniques to randomized trials
  - Improving efficiency using covariates when there is no confounding of effect of interest (randomization works) – Tsiatis, et al. (2006); More generally, Moore and van der laan (2007).
    - If there are predictive covariates, can gain efficiency in estimating the marginal (adjusted) treatment effect, which is protected against model misspecification.

Benefits of Applying Insights of Causal Inference to Clinical Trials

- Can estimate parameters relevant to important questions public health/research questions when randomization does not eliminate confounding.
  - Missing data, informative censoring
  - Isolating pathways (direct effects when intermediates not randomized)

- Provides a general roadmap of estimation (e.g., what is the parameter of interest in the context of competing events)

Examples

- Two applications:
  - the MIRA trial on HIV intervention
  - pain trials and (unmasking) side effects

The MIRA Trial

- Gates Foundation study to determine the effectiveness of a latex diaphragm in the reduction of heterosexual acquisition of HIV among women
- Two arm, randomized, controlled trial
- Primary intervention: diaphragm and gel provision to diaphragm arm (nothing to control arm).
- Secondary Intervention: Intensive condom provision and counseling given to both arms, plus treatment of STIs
- Obviously not blinded.
- 5000 women seen for 18 months in three sites in Zimbabwe and South Africa

MIRA Trial: Basic Intention to Treat Results

- Basic Intent-to-Treat Analysis:
  - 158 new HIV infections in Diaphragm Arm
  - 151 new HIV infections in Control Arm
- ITT estimate of Relative Risk is 1.05 with a 95% CI of (0.84, 1.30)
- End of story . . . . .?

MIRA Trial: Basic Intention to Treat Results

- However condom use differed between the two arms:
  - 53.5% in Diaphragm Arm (by visit)
  - 85.1% in Control Arm (by visit)
- Could this mean that the diaphragm was more effective than it appeared from the basic analysis?
- To make sense of this—we'd like to understand the role of condom use in mediating the effect of treatment assignment on HIV infection.

Most Important Public Health Questions

1. What is the effectiveness of providing study product in environment of country-level standard condom counseling?

   (in environment of no condom counseling?)
2. How does providing study product alone compare to consistent condom use alone in reducing HIV transmission?
3. How does providing the study product alone compare to unprotected sex, in terms of risk of HIV infection?

A Roadmap for Semiparametric Causal Inference

- There is a general framework that can lead, through a series of steps to arrive at:
  1. The relevant parameter of interest (parameter as function of some theoretical intervention)
  2. The identifiability conditions necessary to estimate it (the parameter as a function of the data-generating distribution).
  3. Optimal loss-based estimation of the data-generating distribution.
  4. Defining the locally efficient estimator in a semi-parametric model.
  5. Bending this model so that it targets the parameter of interest.
  6. Derive robust sampling-based inference.

Targeted Learning

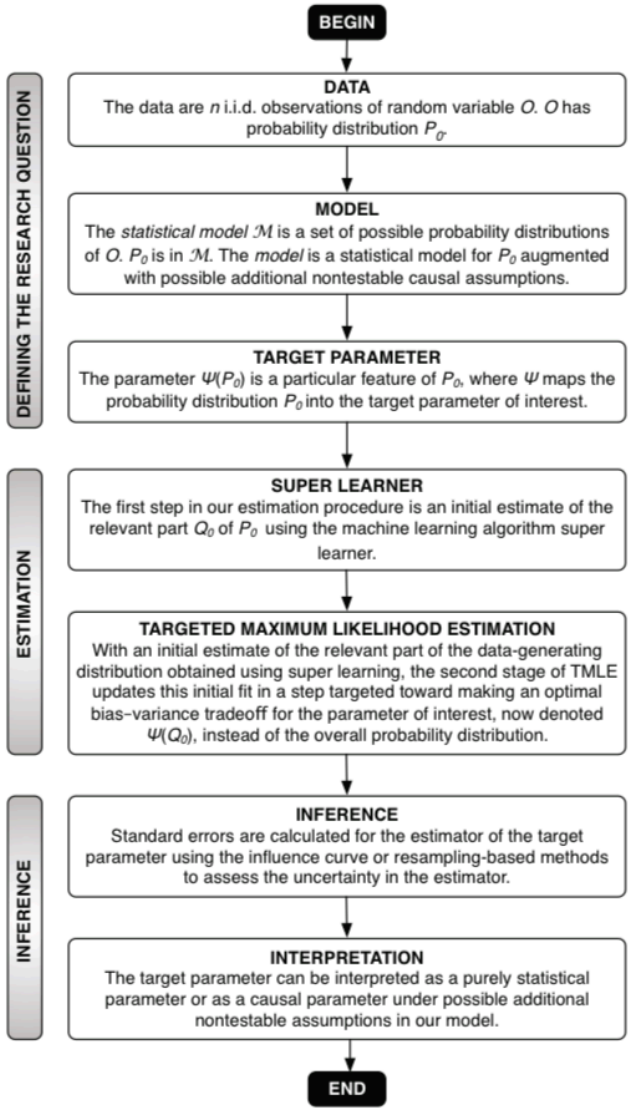
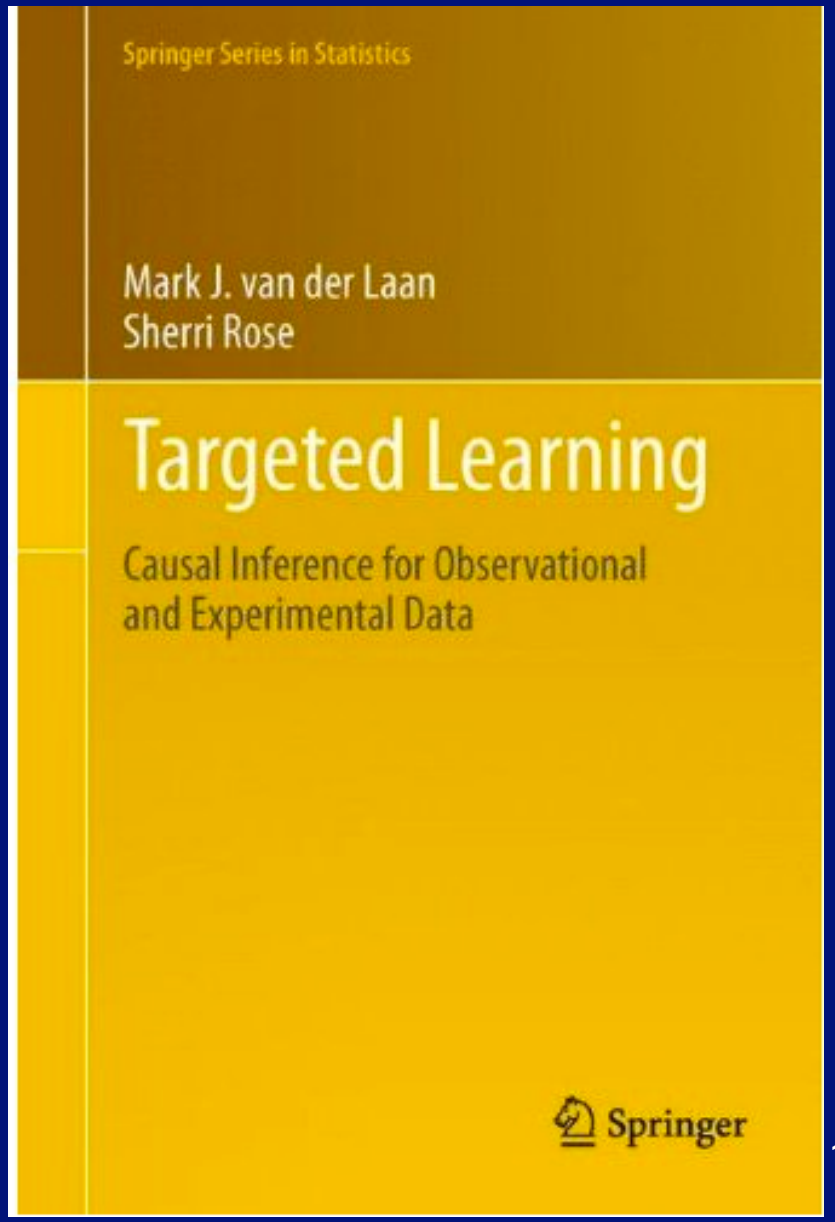


Fig. 1.2 Road map for targeted learning



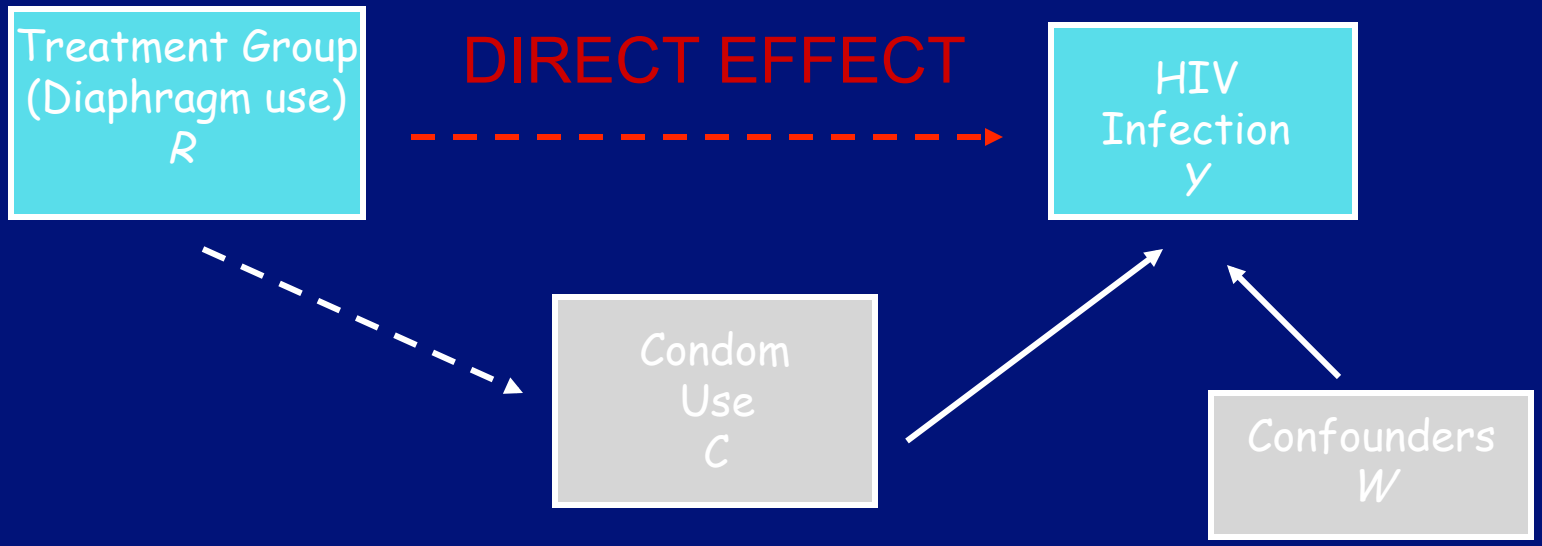
Starting simple

- The actual experiment involved time-dependent measurements of adherence, condom use, HIV status, and baseline covariates.
- To illustrate approach, start with simpler data-structures/questions.
- For instance, single measure of condom use.

The Data, Model 1

- Start with the single definition of condom use case (actual time-dependent measurements).
- $O=(R,W,C,Y)\sim P_0$
  - R=random treatment assignment (Diaphragm=1)
  - W=baseline covariates
  - C=condom use
  - Y=HIV status at end of trial (1=yes)
- Graphical Model: $R=f_R(U_R)$, $W=f_W(U_W)$, $C=f_C(R,U_C)$, $Y=f_Y(R,W,C,U_Y)$, (U's independent)
  - Except for $f_R(U_R)$, which is a known function of a U_R with known distribution, the other functions are unknown, i.e., P_0 nonparametric.?

Graph Associated with Model 1



- We want to estimate the direct effect of diaphragm provision, at a set level of condom use. (Petersen et al. 2006, Robins and Greenland 1992, Pearl 2000, Rosenblum et al. 2009)
- Still ITT interpretation (no confounding of intermediate)

The parameter of interest

- Language:
  - a comparison of the mean if every subject in target or *super* population is assigned diaphragm and uses condoms, versus the same population not assigned diaphragm, but uses condoms.
- Parameter as function of theoretical populations defined by interventions (counterfactual distributions).

$$X = (Y_{rc}, C_r, r \in \mathcal{R}, c \in C), \quad Y_{rc} = f_Y(r, c, U_Y), C_r = f_C(r, U_C)$$

  - Controlled (Type I) DE: $\theta(c) = \psi_c(P_X) = E(Y_{1c} - Y_{0c})$
  - Pure (Type II) DE: $\theta = \psi(P_X) = E(Y_{1C_0} - Y_{0C_0})$

More on Pure Direct Effect

- Can be represented as a weighted average of the controlled direct effects, $\theta(c)$.

$$\theta = \sum_c E(Y_{1c} - Y_{0c})Q(c)$$

$Q(c)=P(C_0=c)$ in this case.

- Given the parameter involves counterfactuals that one never observes, going to need an additional identifiability assumption (, e.g., van der Laan and Petersen, 2004).

$$E(Y_{1c} - Y_{0c} \mid C_0 = c) = E(Y_{1c} - Y_{0c})$$

- Other one's can be invoked (Pearl, 2000; Robins and Greenland, 1992).

More on Pure Direct Effect, cont.

Parameter

- Because need a sort or arbitrary assumption, if one wants a single number (instead of different ones for different *c*'s), then just treat can choose *Q(c)* to be another convenient conditional distribution of *C.*

$$\theta = \sum_{c} E(Y_{1c} - Y_{0c})Q(c)$$

Estimates (Type I)

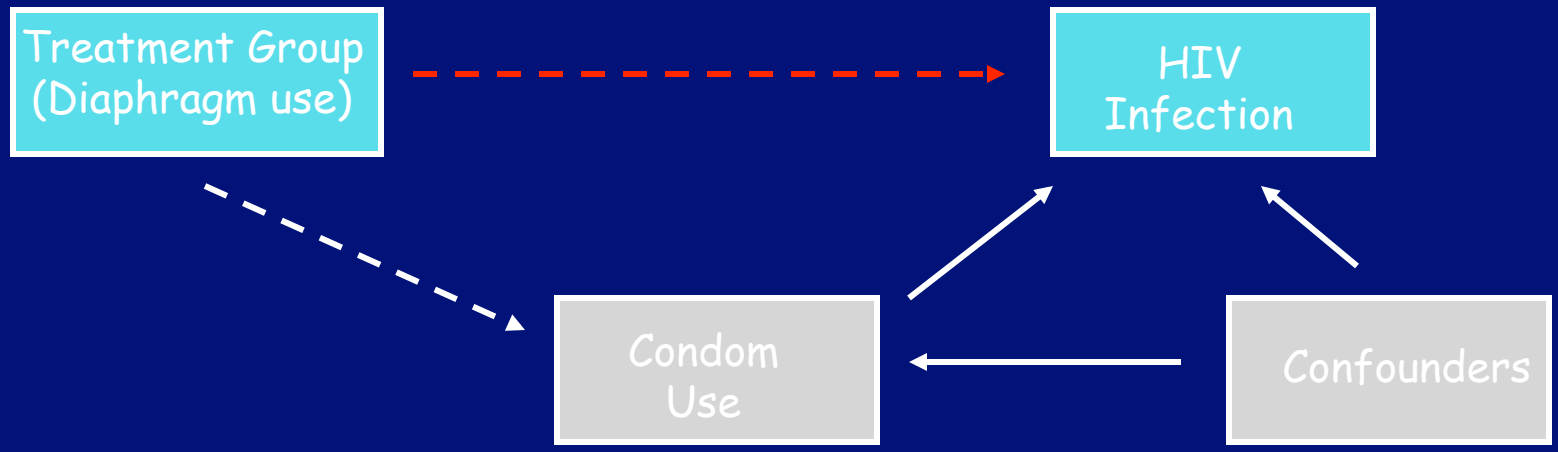
- In this case of controlled direct effect under the very restrictive assumptions of model I, and if we define C at a particular intermediate time then (RA is randomization assumption, CA consistency assump.)

$$E(Y_{1c}) \overset{RA}{=} E(Y_{1c} \mid R=1, C=c) \overset{CA}{=} E(Y \mid R=1, C=c)$$

- Thus, assuming C is discrete (e.g., 1=yes) then a simple plug-in nonparametric estimator.

$$\hat{\theta}(c) = \frac{\sum_{i=1}^{n} Y_i I(R_i=1, C_i=c)}{\sum_{i=1}^{n} I(R_i=1, C_i=c)} - \frac{\sum_{i=1}^{n} Y_i I(R_i=0, C_i=c)}{\sum_{i=1}^{n} I(R_i=0, C_i=c)}$$

Model 2 – Some confounders of condom use



after stratification on condom use

randomization hasn't ruled out confounding of direct effect! How does one define, estimate DE?

The Data, Parameter, RA

- Same data, $O=(R,W,C,Y)\sim P_0$, but now can't ignore W.
- Model: $R=f_R(U_R)$, $W=f_W(U_W)$, $C=f_C(R,W,U_C)$, $Y= f_Y(R,W,C,U_Y)$
- Same parameter of interest:

$$\theta(c) = \psi_c(P_X) = E(Y_{1c} - Y_{0c})$$

- Model implies different randomization assumption:

$$\text{RA} \rightarrow C \perp (Y_{rc}, r \in \mathcal{R}, c \in C) \mid R, W$$

Identifiability

- First, get the $\psi_c(P_0)$:

$$E(Y_{1c} \mid W) \overset{RA}{=} E(Y_{1c} \mid R=1, C=c, W) \overset{CA}{=} E(Y \mid R=1, C=c, W)$$

- But, need to integrate out over *W*:

$$E(Y_{1c}) = \sum_{w} E(Y \mid R=1, C=c, W=w) p(w)$$

Under ETA (positivity): *P(C=c|R=r,W=w) > 0, over dist (R,W).*

$$\underset{\text{want}}{\psi_c(P_X)} \overset{RA/ETA}{=} \underset{\text{can get}}{\psi_c(P_0)} = \sum_{w} \{E(Y \mid R=1, C=c, W=w) - E(Y \mid R=0, C=c, W=w)\} p(w)$$

Estimates

- Two General Approaches
  - Estimating equation (IPCW, DR-IPCW)
  - Plug-in (G-computation, T-MLE)

- DR-IPCW, T-MLE both based on efficient influence curve (asymptotically equivalent)

- Require either estimates of
  - $Q(r,c,W)=E(Y|R=r,C=c,W)$
  - $g(c,r|W)=P(C=c,R=r|W)$
  - both.

- If W high dimensional, requires modeling.

G-computation (Robins)

- Simple plug-in estimator:

$$\hat{\theta}_{Gcomp}(c) = \psi_C(\hat{P}) = \frac{1}{n}\sum_{i=1}^{n}\left\{Q_n^0(1,c,W_i) - Q_n^0(0,c,W_i)\right\}$$

- Need Q^0_n to be consistent for estimate to be consistent.
- Typically know very little about model (semi-parametric)
- Use loss-based approach $\longrightarrow$ Super Learning (van der Laan, Polley, Hubbard, 2007).

Super Learner

CV-risk in Notation

- $\psi_k(R, C, W)$ are candidates for $Q(R, C, W)$ indexed by candidate algorithms (learners), k; $L(Y, \psi_k(\cdot))$ is loss function
- estimate the risk (fit) of each candidate using V-fold X-validation where $B_n(i) = 1$ indicates that for one split, which samples are in validation, remaining observations training set ($Bn(i) = 0$)

$$R_n^{CV}(\hat{\psi}_k, P_n) = \frac{1}{n} \sum_{i:B_n(i)=1} L(Y_i, \psi_k(R_i, C_i, W_i \mid P^0_{n,B_n}))$$

The Empirically Optimal Choice

$$\hat{k} = \underset{k}{\text{argmin}}\ R_n^{CV}(\hat{\psi}_k, P_n)$$

Now, an estimator that minimizes an unbiased estimate of risk. Does it work (give us a theory)?

Super Learner - Optimality

The Oracle - Estimate one would choose if an oracle whispered you the true P_0

Lets assume that there $k = 1, \ldots, K(n)$ algorithms that are competitors. The Oracle one is defines as that which minimizes the

$$\tilde{k}_{B_n,n} = \underset{k}{\operatorname{argmin}} \; E_{B_n} R[\hat{\psi}_k(P^0_{n,B_n}, P_0)]$$

Oracle Inequality

$d_0(\psi, \psi_0) \equiv E_0[L(O, \psi) - L(O, \psi_0)]$, and you need conditions that put bounds on the loss-function (no theorem will deal with outliers). Then, for any $\lambda > 0$:

$$Ed_0[\psi_{\hat{K}(P_n)}(P_{n,T(v)}), \psi_0] \leq$$

$$(1 + 2\lambda) Ed_0[\psi_{\hat{K}(P_n)}(P_{n,T(v)}), \psi_0] + 2C(\lambda) \frac{1 + \log(K(n))}{np}$$

where p is the proportion of the observations in the validation sample, and $C(\lambda)$ is a constant defined in van der Laan et al. (2006).

Bottom Line

- Super learner performs as well as the oracle selector (up to a typically second order term)
- as long as the number of candidate learners considered is polynomial in sample size, no over-fitting.
- can make even better (potentially much better) by weighted averaging over candidates, where this average is determined via cross-validation (so called ensemble learner).

R-function on *CRAN.org*

SuperLearner R function (author Eric Polley)

IPCW (Robins)

- Weighted estimating equation:

$$\hat{E}(Y_{1c}) = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{I(R_i = 1, C = c)Y_i}{g_n(r,c \mid W_i)}\right\}$$

where $g(r,c|W)=P(C=c|W,R)P(R=r|W)=P(C=c|R=r,W)P(R=r)$ (R randomized)

- Based on

$$E\left\{\frac{I(R = r, C = c)Y}{g(r,c \mid W)}\right\} = E_X E\left\{\frac{I(R = r, C = c)Y}{g(r,c \mid W)} \mid X\right\} =$$
$$E_X\left\{\frac{Y}{g(r,c \mid W)} E[I(R = r, C = c)] \mid X\right\} \overset{(R,C)\perp Y_{rc}|W}{=}$$
$$E_X\left\{\frac{Y}{g(r,c \mid W)} g(r,c \mid W)\right\} = EY_{rc}$$

- Need to estimate g (g_n).

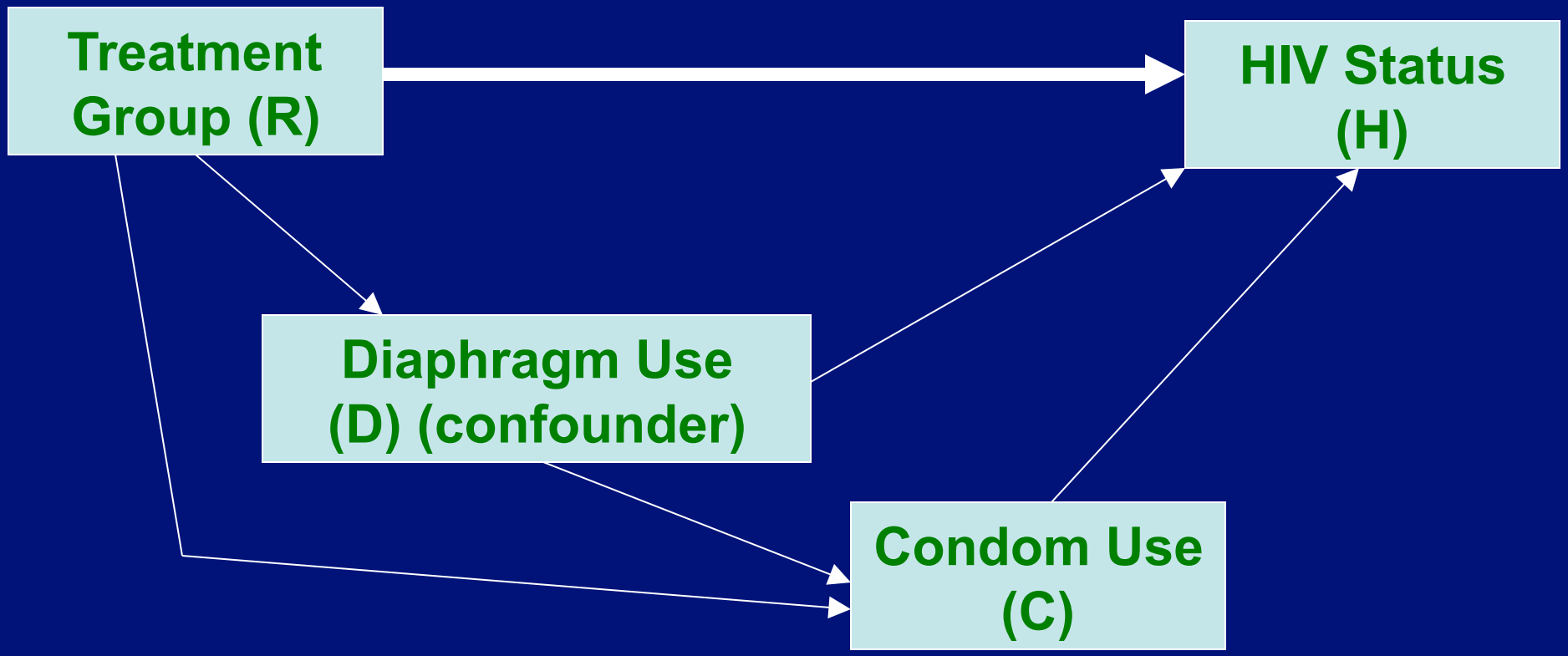
IPTW equivalent can be framed as solving estimating equation

$$m(r,c;\hat{\beta}) : \sum_{i=1}^{n}\left\{\frac{1}{g_n(R_i,C_i \mid W_i)}(Y_i - m(R_i,C_i;\beta))\right\}$$

$$m(r,c;\beta) \equiv \text{saturated model of } E(Y_{rc})$$

- e.g., $m(r,c;\beta) = \beta_0 + \beta_1 r + \beta_2 c + \beta_3 rc$, $E(Y_{1c}-Y_{0c})=\beta_1+\beta_3$.
- Can generalize this to more complicated situations (longitudinal intermediates), and different (non-linear) forms of *m* – called marginal structural models (true of all types of estimators, G-comp, T-MLE, DR-IPTW).

Model 3 (Repeat this for every time, *t = 1,..,8*)



Using Regression, if we control for D, we don't get the direct effect that we want.

Interventions on time-dependent variables

- *O = {W,R,D(0),C(0),Y(0),D(1),C(1),Y(1),…, D(8),C(8),Y(8)}*
- Counterfactuals of interest:

$$Y_{r\bar{c}}(t), \bar{c} = (0,0...,0) \text{ and } \bar{c} = (1,1...,1)$$

- Diaphragm use also an intermediate (can define counterfactuals of it with regards to history of condom use as well).
- To estimate distribution of counterfactuals of interest, need a more general (sequential) randomization assumption

$$C(t) \perp (Y_{r\bar{c}}, D_{r,\bar{c}}) \mid W, R, C(t-1), Y(t-1), D(t-1)$$

Parameter of Interest

Marginal Structural Model

- Lots of counterfactuals of interest – need to make simplifying assumptions (smoothness) – assume a MSM.

- Specifically, model the counterfactual hazard of HIV:

$$\Pr(Y_{r\bar{c}}(t) = 1 \mid Y_{r\bar{c}}(t-1) = 0) = m(t, r, \bar{c}(t) \mid \beta)$$

- Make simplifying (dimension reduction assumptions):
  - condom use measured at the current visit should not have an effect on the outcome of the HIV test at that visit and
  - Condom use in the previous 3 months may have a large effect on whether HIV is detected at a given visit, compared with condom use further in the past.

Estimator - IPTW

- Used inverse weighted estimating equation approach, where weight is proportional to estimated probability of observing the sequential history of condom use for a subject:

$$\prod_{u \le t} \Pr[C(u) \mid W, R, \overline{D}(u), \overline{C}(u-1), \overline{Y}(u-1)]$$

- Sequential randomization assumption (along with even more onerous ETA assumption) and estimating this probability consistently results in consistent estimate of (also needing to be properly specified) hazard model, *m*.

Final Parameter of Interest

- Can use simple product estimator now to get the probability of contracting HIV for different fixed scenarios of condom use and randomization group. If T is time of HIV+, then

$$\hat{P}_r(T_{r\bar{c}} \leq 8) = \prod_{u=1}^{8}\left(1 - m(u,r,\bar{c}(u) \mid \hat{\beta})\right)$$

- Could use any of other methods, and in fact some potentially large advantages to using T-MLE (to appear in *International Journal of Biostatistics*)

- Nonparametric bootstrap used for inference.

Results of Direct Effects Analysis

Results, Rosenblum, et al., 2009

- Relative Risk of HIV infection between Diaphragm arm and Control arm by end of Trial, with Condom Use Fixed at "**Never**": **0.59 (95% CI: 0.26, 4.56)**
- Relative Risk of HIV infection between Diaphragm arm and Control arm by end of Trial, with Condom Use Fixed at "**Always**": **0.96 (95% CI: 0.59, 1.45)**

Conclusion: No definitive evidence from direct effects analysis that diaphragms prevent (or don't prevent) HIV.

Example 2 – Pain Trial

- Pain is the most disturbing symptom of peripheral neuropathy among diabetic patients
- As many as 45% of patients with diabetes develop peripheral neuropathies
- Gabapentin was suggested as a treatment option
- To evaluate the effect of Gabapentin, a randomized, double-blind, placebo-controlled trial was conducted
- 165 patients with a 1- to 5-year history of pain attributed to diabetic neuropathy enrolled at 20 different sites

Backonja et al. Trial in *JAMA*

- The main outcome was daily pain severity as measured on an 11-point Likert scale (0 no pain- 10 worst possible pain)
- Eighty-four patients received gabapentin, 81 received placebo
- By intention-to-treat analysis, gabapentin-treated patients' mean daily pain score (baseline 6.4, end point 3.9) was significantly lower (P<.001) than the placebo-treated patients' score (baseline 6.5, end point 5.1)
- Concluded that gabapentin appears to be efficacious for the treatment of pain associated with diabetic peripheral neuropathy

Handling Treatment-Related Side Effects

- Treat side effects singly by removing those individuals from the data analysis and seeing if that changed the results

Because the study end point of pain was subjective, we explored the possibility that the occurrence of adverse events resulted in unblinding of the study, biasing the result of our efficacy analysis (Table 2). Dizziness and somnolence, the 2 most frequent adverse events, were also those with the largest difference in incidence between the gabapentin and placebo groups. To assess the effect that patients with these events had on the primary efficacy variable we excluded their data and reanalyzed the efficacy data. After excluding data from patients who reported dizziness, the mean pain score between groups differed by –1.19 ($P = .002$), favoring the gabapentin group (gabapentin [n = 62] mean, 4.02; placebo [n = 75] mean, 5.21). After excluding data from patients who reported somnolence, the mean pain score between groups differed by –0.81 ($P = .03$), also favoring the gabapentin group (gabapentin [n = 63] mean, 4.19; placebo [n = 75] mean, 5.21). Thus, inclusion of patients who experienced these central nervous system adverse effects in the original analysis did not account for the overall efficacy seen in the trial.

Perception Effect

- Patients have a perception about the treatment they receive
- In general we may think of the patients assigning a degree of certainty (probability) to receiving the active treatment, measured by a variable *P*

$$P = \begin{cases} 1 & \text{convinced on treatment} \\ 0 & \text{not sure if on treatment or placebo} \\ -1 & \text{convinced on placebo} \end{cases}$$

- In most cases we do not observe the patient's perception on a continuous scale.

Analysis on Side Effects

- Often only the time of occurrence of treatment related side effects
- No equivalent observation on when and if someone might perceive that they are only on placebo (absence of improvement?)
- Previous work (MIRA) indicates issues/ assumptions associated with stratification on side effect occurrence (no longer use data after occurrence of treatment related side effects)

Direct Effects

- Consider an ideal experiment in which the investigator measures the effect of treatment on the outcome holding perception at a fixed level

- **Type I direct effect:** the difference. in the (mean) counterfactual outcomes if the individual received treatment $A = 1$ with her perception fixed at level $P = 0$ vs. the counterfactual outcome if she received no treatment $A = 0$ with her perception fixed at the same level:

$$\psi^{cont}(p) = E(Y_{1p} - Y_{0p})$$

Direct Effects

- **Type II direct effect:**

  the difference in the (mean) counterfactual outcomes if the individual were untreated vs. the counterfactual outcome if she were treated, but her perception remained at its counterfactual level under no treatment:

$$\psi^{Pure} = E(Y_{1P_0} - Y_{0P_0})$$

Data

- *O=(W,A,P, Y)*
- **Outcome:** *(Y)*
  mean pain score for the last 7 diary entries
- **Baseline covariates:** *(W)*
  age, sex, race, height, weight, baseline pain, baseline sleep
- **Treatment:** *(A)*
  gabapentin, placebo
- **Perception:** *(P)*
  changes from 0 to 1 when a treatment-related side effect occurs

Table 1.—Patient Demographics and Baseline Characteristics

Characteristics	Treatment: Gabapentin (n = 84)	Treatment: Placebo (n = 81)
Sex, No. (%)		
Male	49 (58.3)	50 (61.7)
Female	35 (41.7)	31 (38.3)
Race/ethnicity, No. (%)		
White	67 (79.8)	67 (82.7)
Black	5 (6.0)	6 (7.4)
Other	12 (14.3)	8 (9.9)
Age, mean (SD), y	53.0 (10.5)	53.0 (10.2)
Height, mean (SD), cm	173 (13.2)	174 (10.2)
Weight, mean (SD), kg	95.1 (22.6)	94.5 (19.2)
Duration of diabetes, mean (SD), y	12.0 (9.6)	11.2 (8.7)

Parameters of Interest

- What is the treatment effect if all the patients remained unknowledgeable about their treatment? (Perception fixed at 0), $\psi_1=\psi(0)$

preferred ITT parameter

- What is the treatment effect if all the patients thought they were receiving the active treatment? (Perception fixed at 1), $\psi_2=\psi(1)$

- (The difference between these two parameters can be thought of as a *perception bias*)

Parameters of Interest

- What is the perception effect if everyone receives a placebo?(Treatment fixed at 0)

$$\psi_3 = E(Y_{00} - Y_{01}) = E(Y_{00}) - E(Y_{01})$$

- What is the perception effect if everyone receives the active treatment? (Treatment fixed at 1)

$$\psi_4 = E(Y_{10} - Y_{11}) = E(Y_{10}) - E(Y_{11})$$

- (The difference between these parameters yields the same *perception bias*)

Unmasking Bias

- Similarly, the *unmasking bias* can be defined as:

$$\{\mathrm{E}(Y_{11}) - \mathrm{E}(Y_{0,-1})\} - \{\mathrm{E}(Y_{10}) - \mathrm{E}(Y_{00})\}$$

Parameter Estimation Using G-Computation

- Assumptions for G-computation:

  Consistency Assumption:.

  The observed data for a subject is one of the counterfactuals from the full data.

  No Unmeasured Confounding:

  Treatment is randomized within strata of W

  Experimental Treatment Assumption:

$$P(A = a, P = p \mid W) > 0, \ \forall (a,p), P_W$$

- G-comp Estimate

$$E[Y_{ap}] = E_W\big(E(Y|A = a, P = p, W)\big)$$
$$\approx \frac{1}{n}\sum_{i=1}^{n} \hat{E}(Y|A = a, P = p, W_i)$$

$$\mathrm{E}(Y \mid A, P, W)$$

- Estimated using a DSA machine-learning algorithm (forcing in both main effect and interaction terms for *A* and *P,* and up to second degree polynomials in all other terms as needed as determined by 5-fold cross-validation)
- Also used Super Learner

G-comp Estimates

Parameter	Estimate(SE)	P-value	95 % CI
$\psi_1 = E(Y_{00}) - E(Y_{10})$	0.71(0.44)	0.10	(-0.14,1.56)
$\psi_2 = E(Y_{01}) - E(Y_{11})$	2.50(0.69)	0.0002	(1.15,3.84)
$\psi_3 = E(Y_{00}) - E(Y_{01})$	-0.59(0.64)	0.35	(-1.83,0.65)
$\psi_4 = E(Y_{10}) - E(Y_{11})$	1.18(0.51)	0.02	(0.18,2.17)

Parameter	Estimate(SE)
$E(Y_{00})$	5.14(0.25)
$E(Y_{01})$	5.74(0.60)
$E(Y_{10})$	4.42(0.37)
$E(Y_{11})$	3.24(0.37)

Alternatives to G-Computation

- Inverse probability (of “treatment”) weighting--probably less efficient
- Double-robust version of IPTW—needs specialized software
- Targeted Maximum Likelihood Est (TMLE) extension of G-computation (and asymptotically equivalent to the double-robust estimator)—allows use of standard software
- Collaborative TMLE

Targeted Maximum Likelihood Estimation

- Observed data: $O = (W, A, P, Y) \sim P_0$
- Suppose we want to estimate $\hat{E}[Y_{10}]$
- Given the assumptions the likelihood can be written as:

$$P_0(O) = P_0(W)P_0(A = a, P = p|W)P_0(Y|A = a, P = p, W)$$

$$= P_0(W)P_0(A = a)P_0(P = p|A = a, W)\ P_0(Y|A = a, P = p, W)$$

$$\hat{E}[Y_{10}] = \frac{1}{n}\sum_{i=1}^{n} \hat{Q}_0(1,0,W) + \frac{\hat{\varepsilon}_{10}}{g_{10}(A = 1, P = 0|W)}$$

TMLE Estimates

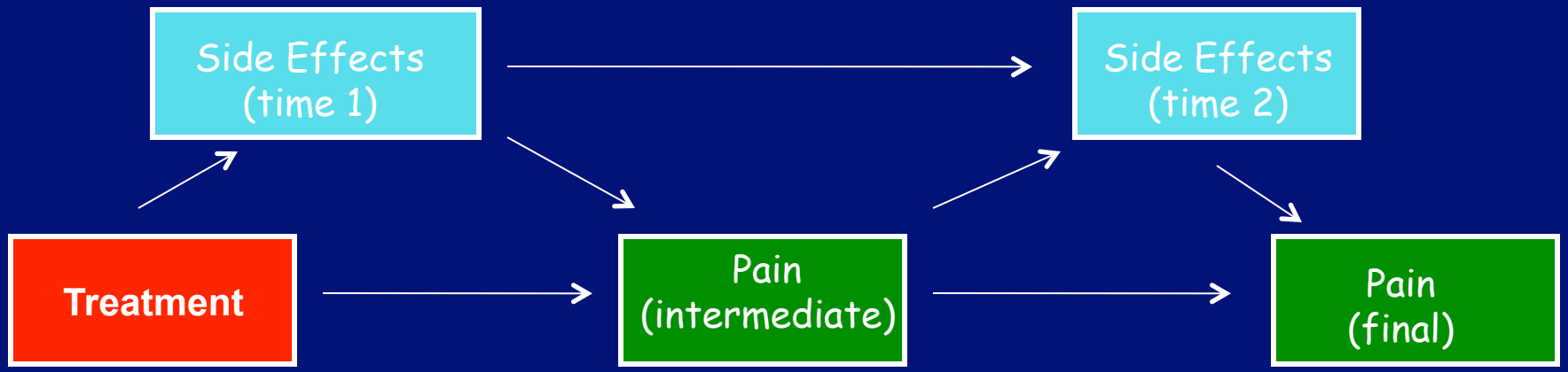
Parameter	Estimate(SE)	P-value	95 % CI
$\psi_1 = E(Y_{00}) - E(Y_{10})$	0.78(0.58)	0.18	(-0.35,1.91)
$\psi_2 = E(Y_{01}) - E(Y_{11})$	1.98(0.99)	0.04	(0.04,3.91)
$\psi_3 = E(Y_{00}) - E(Y_{01})$	-0.07(0.81)	0.93	(-1.64,1.50)
$\psi_4 = E(Y_{10}) - E(Y_{11})$	1.12(0.62)	0.07	(-0.09,2.32)

Parameter	Estimate(SE)
$E(Y_{00})$	5.15(0.31)
$E(Y_{01})$	5.22(0.77)
$E(Y_{10})$	4.35(0.41)
$E(Y_{11})$	3.39(0.47)

Potential Problems/Extensions

- ETA Bias: Get small probability estimates for $g(P,A|W)$, inflated standard errors
- Use full longitudinal feature of the data (daily recorded pain scores—apparently makes little difference to what has been shown here)

Time Dependent Intermediaries



- Time dependent confounding if the intermediate pain scores are ignored

Implications for Design/Analysis?

- Causal methods help us think about what we want to estimate and appropriate methods to collect data to achieve this goal within an honest model (usually nonparametric).
- Ethics of intensive condom counseling--human subjects review?
- Alternative (adaptive) designs (focus on non-condom users, adherents etc)
- How do we measure intermediate variables (eg condom use, side effects) effectively?
- Need to think about measurement of potential confounders even with randomization?
- Use of surrogate outcomes (eg HSV?, objective pain measurements?) and comparison with outcomes of interest
- Measurement of perception for all subjects in RCTs with self-reported outcomes