Post-Randomization Factors, Direct Effects and Randomized Trials

Nicholas P. Jewell
Department of Statistics &
School of Public Health (Biostatistics)
University of California, Berkeley

Recent Advances in Statistics for Causal Analysis
June 8, 2011, Bordeaux

Collaborators

Alan Hubbard

Mark van der Laan

Farid Jamshidian

Michael Rosenblum

Supported by NIAID

References

- 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

Purpose of Talk

- The application of causal inference techniques to randomized trials
  - the basic Intent-to-Treat (ITT) analysis may not answer the most important public health/research questions
  - Why are causal methods needed?
  - What questions do they answer?
  - Are these the right questions?
- We present ideas, requiring more assumptions than the ITT, for answering some of these questions.
- Three major applications
  - the MIRA trial on HIV intervention
  - pain trials and (unmasking) side effects
  - The use of statins in cardiovascular safety studies

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
- Trial is 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.

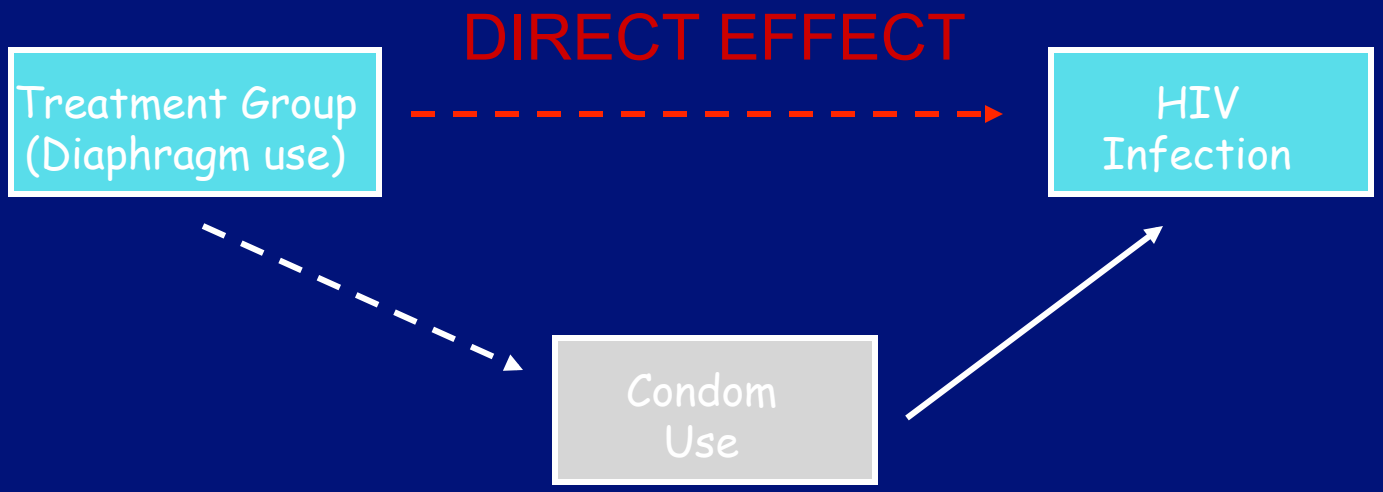
Most Important Public Health Questions

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?

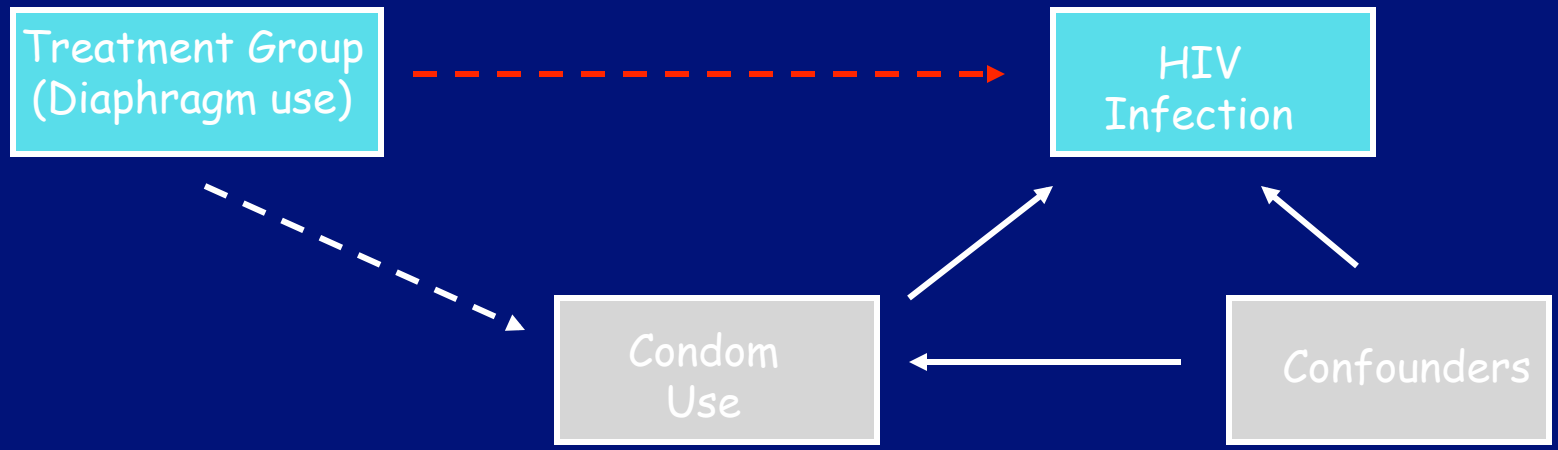
None of these questions are answered by basic ITT analysis

Estimating Direct Effects: Adjusting for a Mediator (condom use)

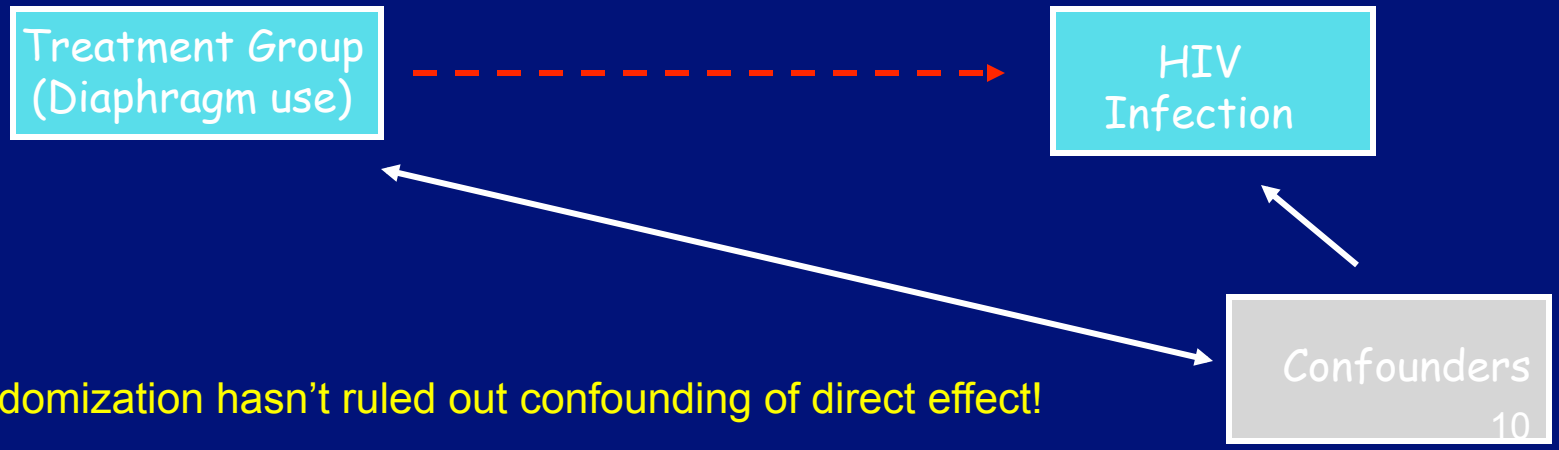


- We want to estimate the direct effect of diaphragm provision, at a set level of condom use. (Petersen et al. 2006, Robins and Greenland 1992, Pearl 2000, Rosenblum et al. 2009)
- Still ITT interpretation
- Requires Stronger Assumptions than basic ITT

Estimating Direct Effects



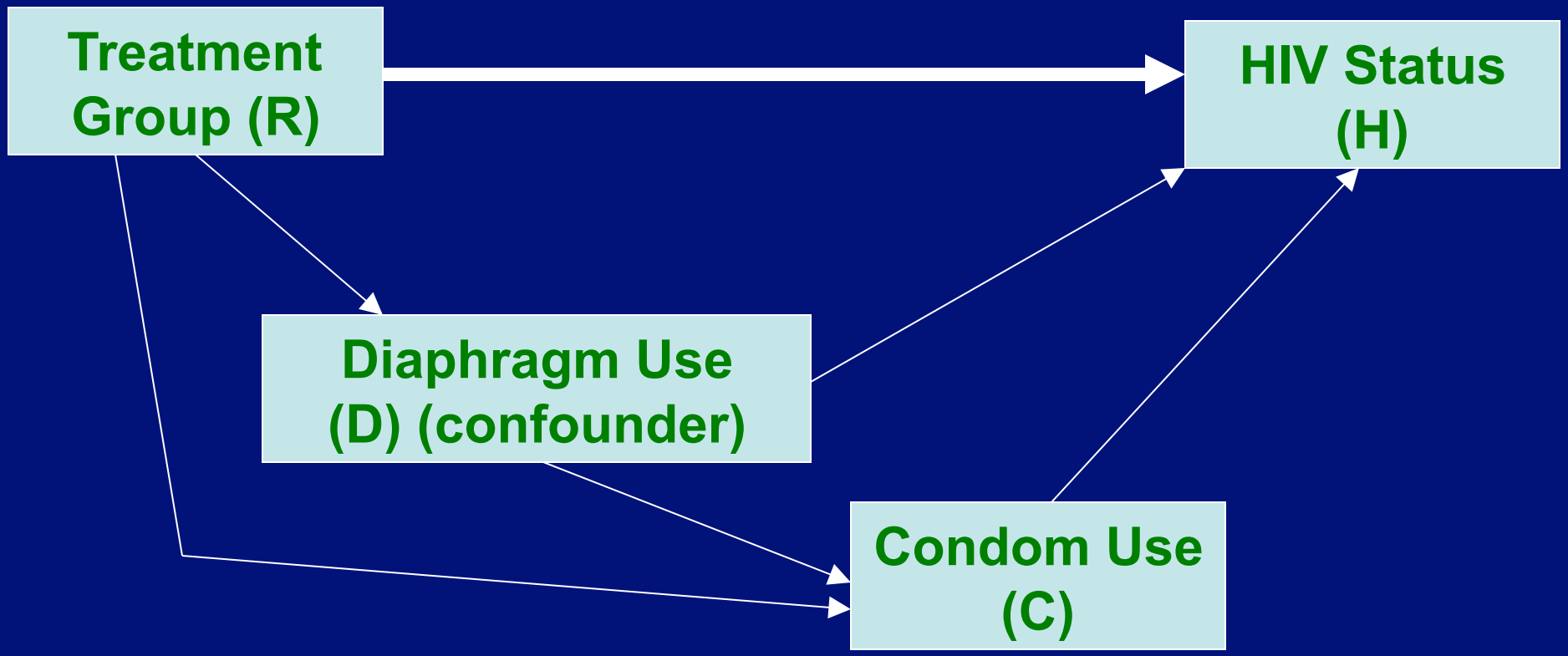
after stratification on condom use



randomization hasn't ruled out confounding of direct effect!

Now have to adjust for confounders (but we are still ITT)

Why Stratification/Regression Gives Biased Estimates When There Are Confounders as Causal Intermediates



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

Results of Direct Effects Analysis

- 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.

Pain Trials with Self-Reporting

- 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

Data/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:

$$Y_{1p} - Y_{0p}$$

Data

- **Outcome:**
  mean pain score for the last 7 diary entries
- **Baseline covariates:**
  age, sex, race, height, weight, baseline pain, baseline sleep
- **Treatment:**
  gabapentin, placebo
- **Perception:**
  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)

$$\psi_1 = E(Y_{00} - Y_{10}) = E(Y_{00}) - E(Y_{10})$$

preferred ITT parameter

- What is the treatment effect if all the patients thought they were receiving the active treatment? (Perception fixed at 1)

$$\psi_2 = E(Y_{01} - Y_{11}) = E(Y_{01}) - E(Y_{11})$$

- (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:

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

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:

$$P(A = a, P = p|W) > 0 \quad W\forall$$

- Estimate by

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

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

- 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)

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

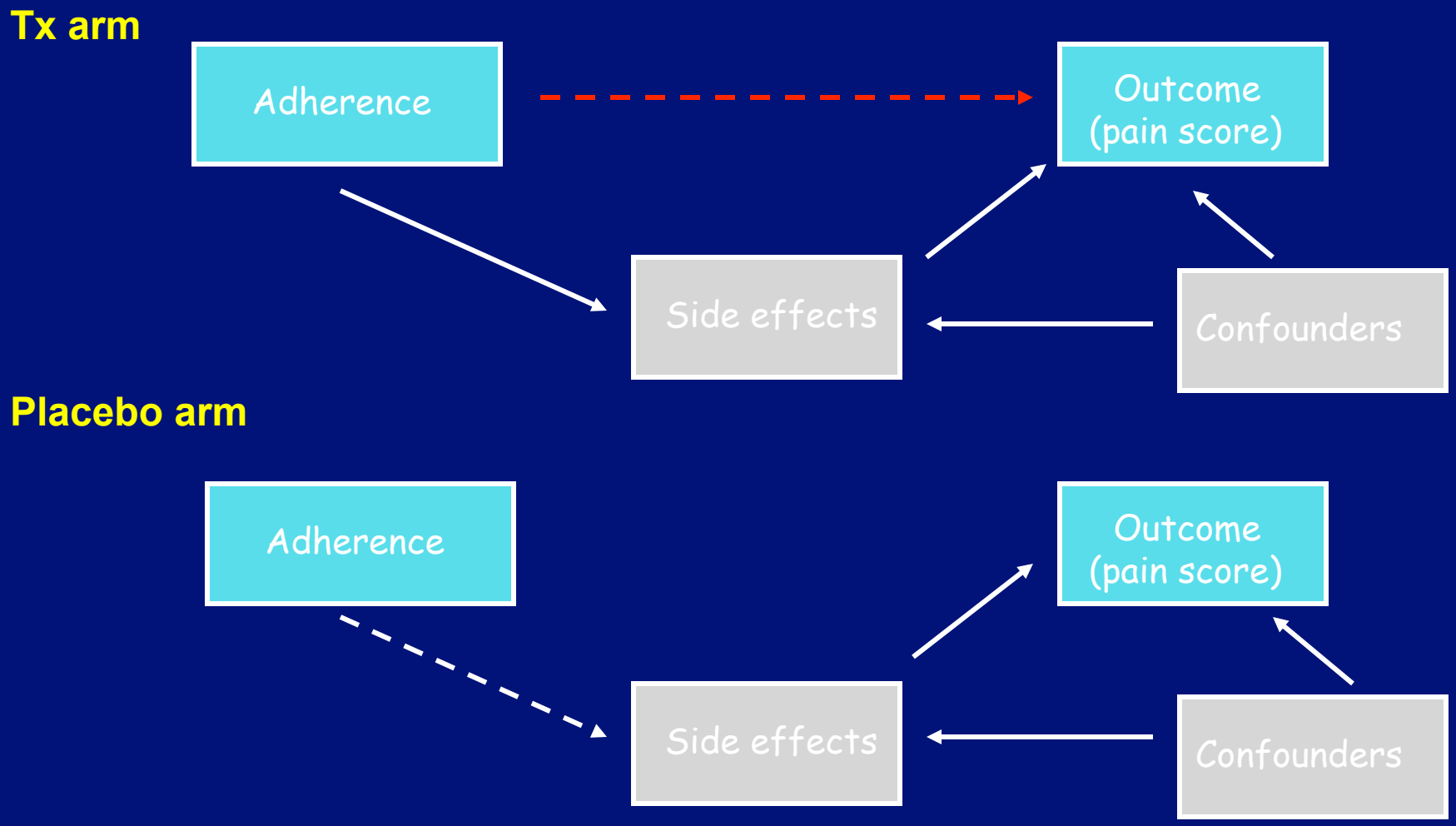
- Inverse probability (of “treatment”) weighting--probably less efficient
- Double-robust version of IPTW—needs specialized software
- Targeted Maximum Likelihood (TML) extension of G-computation (and asymptotically equivalent to the double-robust estimator)—allows use of standard software

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)

Advantages/Differences of Blinded Trials

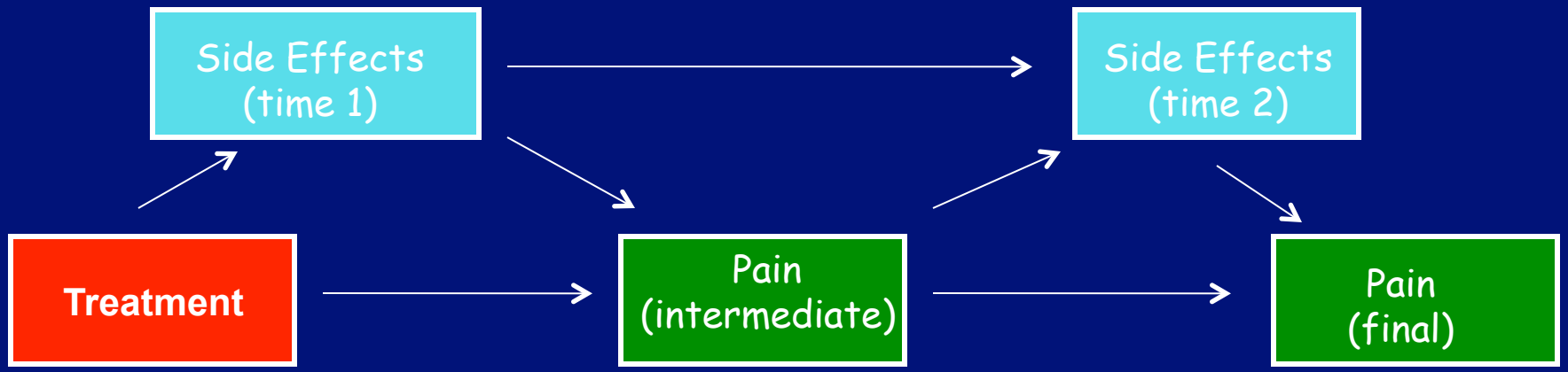


- Two different causal diagrams for the effect of adherence depending on treatment arm (can this be exploited, other than through use of instrumental variable approach?)

Advantages/Differences of Blinded Trials

- Different causal diagrams in each arm
- Testing the null is appropriate even with differential rate of side effects amongst adherents/non-adherents
- ITT analysis will not ‘reverse’ effects

Time Dependent Intermediaries



- Time dependent confounding if the intermediate pain scores are ignored

Statin Use in Cardiovascular Safety Studies

- The FIELD (Fenofibrate Intervention and Event Lowering in Diabetes) study diagrams in each arm: increased use of statins in placebo group (Simes et al., 2010)
- The RECORD (Rosiglitazone Evaluated for Cardiac Outcomes and Regulation of glycemia in Diabetes) trial examining the cardiovascular safety of Avandia—ope label with more statins for the Avandia group.

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
- 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 in MIRA?, objective pain measurements?) and comparison with outcomes of interest
- Measurement of perception for all subjects in RCTs with self-reported outcomes