

Evaluating Candidate Surrogate Endpoints with Principal Stratification

Cory Zigler

CHIPTS Methods Seminar

December 2, 2008

Outline

Introduction

- Potential Outcomes
- Posttreatment Variables
- Principal Stratification

Surrogate Endpoints

- Definition
- Evaluating Surrogates

Estimating Surrogate Value

- Causal Effect Predictiveness (CEP)
- Illustration

Outline

Introduction

Potential Outcomes

Posttreatment Variables

Principal Stratification

Surrogate Endpoints

Definition

Evaluating Surrogates

Estimating Surrogate Value

Causal Effect Predictiveness (CEP)

Illustration

One-Person Clinical Trial

HIV therapy

- ▶ give 1 HIV+ person a new therapy ($Z_i = 1$)
- ▶ follow up for 5 years
- ▶ observe that the person does not contract any serious opportunistic infections (OIs), ($Y_i = 0$)
- ▶ did the therapy *cause* the person to not have an OI?

Cause and Effect?

Would this person have had an OI with the standard (control) therapy?

- ▶ could compare what *did* happen to what *would have* happened
 - ▶ OI free with new therapy + OI with control therapy \Rightarrow new therapy prevented OI
- ▶ only get to observe one outcome

Potential Outcomes Notation

Potential outcome under both possible treatments, $Y(Z)$

- ▶ $Y_i(1)$: outcome if $Z_i = 1$ (new)
- ▶ $Y_i(0)$: outcome if $Z_i = 0$ (control)
- ▶ person has both
- ▶ assume assignment to Z determines which $Y(Z)$ becomes observed, but does not affect $Y(0)$ or $Y(1)$
- ▶ innate characteristics that exist before application of treatment

Causal Effect

The causal effect of Z on Y for the i^{th} patient is defined by

$$Y_i(1) - Y_i(0)$$

Fundamental problem: we only observe either $Y_i(0)$ or $Y_i(1)$

- ▶ could look for another person that is *exactly* the same as the i^{th} person, and give that person the other treatment
- ▶ could give the i^{th} person new treatment *then* give the standard treatment (temporal stability)

Randomized Trial

Randomize 1000 patients to receive new ($Z = 1$) or control ($Z = 0$) treatment

- ▶ no serious OIs in the treatment arm
- ▶ 30% of the patients in the control arm get a serious OI
- ▶ randomization \Rightarrow treatment arms have same distribution of background characteristics
- ▶ observed control outcomes represent what *would have happened* to the new treatment recipients
 - ▶ if units are assumed exchangeable
- ▶ conclude that the new treatment *causes* a person to remain OI-free

Average Causal Effects

If treatment groups are comparable (eg, randomized study)

$$\bar{Y}(1) - \bar{Y}(0)$$

If treatment groups are *not* comparable (eg, observational study)

- ▶ observed differences
 - ▶ stratify subjects on covariates so that units are comparable within strata, then assess $\bar{Y}(1) - \bar{Y}(0)$
- ▶ unobserved differences
 - ▶ can undermine foundation for causal inference

Outline

Introduction

Potential Outcomes

Posttreatment Variables

Principal Stratification

Surrogate Endpoints

Definition

Evaluating Surrogates

Estimating Surrogate Value

Causal Effect Predictiveness (CEP)

Illustration

Posttreatment Variable

Randomized clinical trial

- ▶ treatment ($Z = 1$) vs. control ($Z = 0$)
- ▶ posttreatment variable, S
 - ▶ measured *after* randomization

$S(Z)$ is an outcome (depends on treatment)

- ▶ potential posttreatment outcomes: $S(0)$, $S(1)$
- ▶ causal effect of Z on S : $S(1) - S(0)$

Adjusting for a Posttreatment Variable

“Adjusting” for S

- ▶ Z affects $S \Rightarrow$ treatment groups imbalanced on S
 - ▶ $S(1)$ observed in treatment arm
 - ▶ $S(0)$ observed in control arm
- ▶ S affects $Y \Rightarrow S$ imbalance confounds treatment effect on Y
- ▶ S is *concomitant* (associated with outcome and treatment)
 - ▶ Cochran (*Biometrics*, 1957): In ANCOVA, avoid adjusting for posttreatment concomitant variable
- ▶ adjusting for observed S can “adjust away” the treatment effect

Example: HIV Therapy Trial with Posttreatment Variable

Randomly assign $n=1000$ patients

- ▶ treatment: control ($Z = 0$) vs. new ($Z = 1$) therapy
- ▶ outcome: serious OI ($Y = 1$) or OI-free ($Y = 0$) at 5 year follow-up
- ▶ S: low (L) or high (H) CD4 count measured at 1 year follow-up
- ▶ successful treatment \Rightarrow high CD4 \Rightarrow smaller chance of OI

HIV Therapy Trial Underlying Data

S(0)	S(1)	n	Y(0)	Y(1)
L	L	500	380 (76%)	440 (88%)
L	H	300	150 (50%)	100 (33%)
H	H	200	60 (30%)	60 (30%)

HIV Therapy Trial Observed Data

Z	S(0)	S(1)	n	Y(0)	Y(1)
0	L	L	250	190 (76%)	220 (88%)
0	L	H	150	75 (50%)	50 (33%)
0	H	H	100	30 (30%)	30 (30%)
1	L	L	250	190 (76%)	220 (88%)
1	L	H	150	75 (50%)	50 (33%)
1	H	H	100	30 (30%)	30 (30%)

$$\bar{Y}(0) = .59$$

$$\bar{Y}(1) = .6$$

Overall “intent-to-treat” treatment effect based on observed data:

$$.59 - .6 = -.01$$

HIV Therapy Trial Observed Data

Z	S(0)	S(1)	n	Y(0)	Y(1)
0	L	L	250	190 (76%)	220 (88%)
0	L	H	150	75 (50%)	50 (33%)
0	H	H	100	30 (30%)	30 (30%)
1	L	L	250	190 (76%)	220 (88%)
1	L	H	150	75 (50%)	50 (33%)
1	H	H	100	30 (30%)	30 (30%)

$$\bar{Y}(0) = .59$$

$$\bar{Y}(1) = .6$$

Overall “intent-to-treat” treatment effect based on observed data:

$$.59 - .6 = -.01$$

HIV Therapy Trial Observed Data

Z	S(0)	S(1)	n	Y(0)	Y(1)
0	L	L	250	190 (76%)	220 (88%)
0	L	H	150	75 (50%)	50 (33%)
0	H	H	100	30 (30%)	30 (30%)
1	L	L	250	190 (76%)	220 (88%)
1	L	H	150	75 (50%)	50 (33%)
1	H	H	100	30 (30%)	30 (30%)

$$\bar{Y}(0) = .59$$

$$\bar{Y}(1) = .6$$

Overall “intent-to-treat” treatment effect based on observed data:

$$.59 - .6 = -.01$$

HIV Therapy Trial Observed Data

Z	S(0)	S(1)	n	Y(0)	Y(1)
0	L	L	250	190 (76%)	220 (88%)
0	L	H	150	75 (50%)	50 (33%)
0	H	H	100	30 (30%)	30 (30%)
1	L	L	250	190 (76%)	220 (88%)
1	L	H	150	75 (50%)	50 (33%)
1	H	H	100	30 (30%)	30 (30%)

$$\bar{Y}(0) = .59$$

$$\bar{Y}(1) = .6$$

Overall “intent-to-treat” treatment effect based on observed data:

$$.59 - .6 = -.01$$

HIV Therapy Trial Data: Adjust for CD4

Z	S(0)	S(1)	n	Y(0)	Y(1)
0	L	L	250	190 (76%)	220 (88%)
0	L	H	150	75 (50%)	50 (22%)
0	H	H	100	30 (30%)	30 (30%)
1	L	L	250	190 (76%)	220 (88%)
1	L	H	150	75 (50%)	50 (33%)
1	H	H	100	30 (30%)	30 (30%)

$$\bar{Y}(0)|L = .662 \quad \bar{Y}(1)|L = .88$$

Observed effect: -.218

$$\bar{Y}(0)|H = .30 \quad \bar{Y}(1)|H = .32$$

Observed effect: -.02

HIV Therapy Trial Data: Adjust for CD4

Z	S(0)	S(1)	n	Y(0)	Y(1)
0	L	L	250	190 (76%)	220 (88%)
0	L	H	150	75 (50%)	50 (22%)
0	H	H	100	30 (30%)	30 (30%)
1	L	L	250	190 (76%)	220 (88%)
1	L	H	150	75 (50%)	50 (33%)
1	H	H	100	30 (30%)	30 (30%)

$$\bar{Y}(0)|L = .662 \quad \bar{Y}(1)|L = .88$$

Observed effect: -.218

$$\bar{Y}(0)|H = .30 \quad \bar{Y}(1)|H = .32$$

Observed effect: -.02

HIV Therapy Trial Data: Adjust for CD4

Z	S(0)	S(1)	n	Y(0)	Y(1)
0	L	L	250	190 (76%)	220 (88%)
0	L	H	150	75 (50%)	50 (22%)
0	H	H	100	30 (30%)	30 (30%)
1	L	L	250	190 (76%)	220 (88%)
1	L	H	150	75 (50%)	50 (33%)
1	H	H	100	30 (30%)	30 (30%)

$$\bar{Y}(0)|L = .662 \quad \bar{Y}(1)|L = .88$$

Observed effect: -.218

$$\bar{Y}(0)|H = .30 \quad \bar{Y}(1)|H = .32$$

Observed effect: -.02

Adjusting for S

New treatment \Rightarrow more OIs?

- ▶ treatment appears to cause high CD4, on average
- ▶ adjusting for post-treatment CD4 “adjusts away” this effect
- ▶ result is the **net effect** of
 - ▶ treatment
 - ▶ unobserved characteristics that lead to different values of S under treatment and control
- ▶ treatment/control groups not comparable within observed level of S
- ▶ posttreatment selection bias

Outline

Introduction

Potential Outcomes

Posttreatment Variables

Principal Stratification

Surrogate Endpoints

Definition

Evaluating Surrogates

Estimating Surrogate Value

Causal Effect Predictiveness (CEP)

Illustration

Principal Stratification (Frangakis and Rubin (*Biometrics*, 2002))

Principal Strata, $S_i = (S_i(0), S_i(1))$

- ▶ $(S_i(0) = L, S_i(1) = L) \equiv$ “weaker” immune systems
- ▶ $(S_i(0) = L, S_i(1) = H) \equiv$ “normal” immune systems
- ▶ $(S_i(0) = H, S_i(1) = H) \equiv$ “stronger” immune systems

principal strata are *not* affected by treatment (i.e., play a role akin to covariates)

HIV Therapy Trial: Principal Effects

Compare $\bar{Y}(0) - \bar{Y}(1)$ *within strata*

- ▶ weaker: $\bar{Y}(0) - \bar{Y}(1) = .76 - .88 = -.12$
- ▶ normal: $\bar{Y}(0) - \bar{Y}(1) = .5 - .33 = .17$
- ▶ stronger: $\bar{Y}(0) - \bar{Y}(1) = .3 - .3 = 0$

Conclusion: treatment effect depends on CD4 response

- ▶ treatment hurts “weaker” patients
- ▶ treatment helps “normal” patients
- ▶ treatment does not affect “stronger” patients

Principal Effects

Comparisons within principal strata

- ▶ account for potential values of S
- ▶ control for unobserved characteristics that lead to certain values of S
 - ▶ e.g., patients with weaker immune systems are more likely to get OIs
- ▶ are comparisons between comparable groups of patients
- ▶ yield causal interpretations

Outline

Introduction

- Potential Outcomes
- Posttreatment Variables
- Principal Stratification

Surrogate Endpoints

- Definition**
- Evaluating Surrogates

Estimating Surrogate Value

- Causal Effect Predictiveness (CEP)
- Illustration

Surrogate Endpoints (Biomarkers) in Clinical Trials

Definition

A **posttreatment** value that reliably predicts correct inferences regarding the effect of a treatment on a clinical endpoint

Uses

- ▶ avoid long follow up
- ▶ reduce costs
- ▶ replace difficultly obtained clinical outcomes
- ▶ elucidate causal mechanisms

Example: HIV Field Study

Randomly assign $n=1000$ patients

- ▶ Z : intervention ($Z = 1$) or control ($Z = 0$)
- ▶ Y : PTSD ($Y = 1$) or no PTSD ($Y = 0$)
- ▶ S : stress biomarker from saliva
 - ▶ salivary alpha-amylase (SAA)
- ▶ X : pre-treatment covariates

Rationale/Goal

- ▶ might be expensive/difficult/impossible to translate PTSD instruments into native language
- ▶ evaluate if SAA can be serve as a valid surrogate endpoint

Example: HIV Field Study

Randomly assign $n=1000$ patients

- ▶ Z : intervention ($Z = 1$) or control ($Z = 0$)
- ▶ Y : PTSD ($Y = 1$) or no PTSD ($Y = 0$)
- ▶ S : stress biomarker from saliva
 - ▶ salivary alpha-amylase (SAA)
- ▶ X : pre-treatment covariates

Rationale/Goal

- ▶ might be expensive/difficult/impossible to translate PTSD instruments into native language
- ▶ evaluate if SAA can serve as a valid surrogate endpoint

What Makes an Ideal Surrogate?

Intervention effect on SAA should reliably predict effect on PTSD outcomes

- ▶ Z causally affects S
 - ▶ intervention induces change in SAA
- ▶ S causally affects Y
 - ▶ SAA induces change in chance of PTSD
- ▶ Z only affects Y *through* S
 - ▶ intervention does not affect PTSD through any other “causal pathway”

Outline

Introduction

- Potential Outcomes
- Posttreatment Variables
- Principal Stratification

Surrogate Endpoints

- Definition
- Evaluating Surrogates**

Estimating Surrogate Value

- Causal Effect Predictiveness (CEP)
- Illustration

Principal Effects: Compare Outcomes Within Strata ($S(0)$, $S(1)$)

Probability of PTSD under **intervention** in
($S(0) = s_0$, $S(1) = s_1$)

$$\text{risk}_{(1)}(s_0, s_1) \equiv \Pr(Y(1) = 1 | S(0) = s_0, S(1) = s_1)$$

Probability of PTSD under **control** in
($S(0) = s_0$, $S(1) = s_1$)

$$\text{risk}_{(0)}(s_0, s_1) \equiv \Pr(Y(0) = 1 | S(0) = s_0, S(1) = s_1)$$

Criteria for Evaluating Surrogate Endpoints

Surrogate questions

- ▶ does the intervention affect PTSD outcomes through means other than SAA?
 - ▶ can the intervention reduce PTSD without affecting SAA?
- ▶ how much of the intervention effect is “captured” by SAA?
 - ▶ does a change in SAA *cause* a change in PTSD?
 - ▶ does a larger effect on SAA correspond to a larger effect on the chance of PTSD?

Surrogate Question 1: Does the intervention affect PTSD outcomes through means other than SAA?

For subjects with no effect on SAA

- ▶ $(S(0) = s_0, S(1) = s_1)$ with $s_1 = s_0$
- ▶ compare $risk_{(1)}(s_0, s_1)$ vs. $risk_{(0)}(s_0, s_1)$
- ▶ $risk_{(1)}(s_0, s_1) \neq risk_{(0)}(s_0, s_1) \Rightarrow$ intervention affects PTSD without changing SAA
- ▶ $risk_{(1)}(s_0, s_1) = risk_{(0)}(s_0, s_1) \Rightarrow$ intervention does not affect PTSD without changing SAA

Surrogate Question 1: Does the intervention affect PTSD outcomes through means other than SAA?

Dissociative effect

- ▶ residual causal effect of intervention that occurs without changing SAA
- ▶ nonzero dissociative effect \Rightarrow additional causal pathway
 - ▶ causal pathway that doesn't "pass through" S
- ▶ no dissociative effect \Rightarrow *causal necessity*
 - ▶ a change in S is necessary for a change in Y

Surrogate Question 2: How much of the intervention effect is “captured” by SAA?

For patients with an effect on SAA

- ▶ $(S(0) = s_0, S(1) = s_1)$, with $s_1 \neq s_0$
- ▶ compare $risk_{(1)}(s_1, s_0)$ vs. $risk_{(0)}(s_1, s_0)$
- ▶ $risk_{(1)}(s_0, s_1) \neq risk_{(0)}(s_0, s_1) \Rightarrow$ intervention affects PTSD outcomes *through* SAA
- ▶ check if larger $|s_1 - s_0|$ indicates larger risk difference

Surrogate Question 2: How much of the intervention effect is “captured” by SAA?

Associative effect

- ▶ intervention effect *associated* with inducing a change in SAA from s_0 to s_1
- ▶ causal pathway between Z and Y that “passes through” S
- ▶ nonzero associative effect \Rightarrow *causal sufficiency*
 - ▶ for some $|s_1 - s_0| > C$, $risk_{(1)}(s_0, s_1) \neq risk_{(0)}(s_0, s_1)$

Principal Surrogate

A **principal surrogate** is one that satisfies causal necessity and causal sufficiency

- ▶ a change in Y *requires* a change in S
- ▶ a change in S induces a change in Y
- ▶ requires knowledge both $S(0)$ and $S(1)$
- ▶ compare with standard definitions of valid surrogates based on observed S (*statistical surrogates*)
 - ▶ Prentice Criteria (Prentice (*Statistics in Medicine*, 1989))

Surrogate Value

Principal Surrogates

- ▶ causal necessity \nRightarrow valuable surrogate
- ▶ valuable surrogate \nRightarrow causal necessity

Associative effect relative to dissociative effect

- ▶ large associative + small dissociative \Rightarrow valuable surrogate
- ▶ a large proportion of the overall ‘intent-to-treat’ intervention effect is associated with the biomarker

Outline

Introduction

- Potential Outcomes
- Posttreatment Variables
- Principal Stratification

Surrogate Endpoints

- Definition
- Evaluating Surrogates

Estimating Surrogate Value

- Causal Effect Predictiveness (CEP)
- Illustration

CEP Surface (Gilbert et al (*Biometrics*, 2008))

Contrast function $h(x, y)$

- ▶ $h(x, y) = 0$ if $x = y$
- ▶ eg, $x - y$, $\log(x/y)$

$$CEP^{risk}(s_1, s_0) \equiv h(risk_{(1)}(s_0, s_1), risk_{(0)}(s_0, s_1))$$

- ▶ principal effects
- ▶ $CEP^{risk}(s, s) = 0$ for all $s \Rightarrow$ causal necessity
- ▶ $CEP^{risk}(s_1, s_0) \neq 0$ for all $|s_1 - s_0| > C \Rightarrow$ causal sufficiency

Missing Potential Outcomes \Rightarrow CEP Surface Not Identified

Impute missing $S(1 - Z) \Rightarrow$ identify CEP surface (current strategies)

- ▶ assume constant biomarker in control group
- ▶ covariate(s) predictive of $S(0)$ or $S(1)$
 - ▶ irrelevant vaccine (Follmann, (*Biometrics*, 2006))
- ▶ “closeout” trials (Follmann, (*Biometrics*, 2006))
- ▶ estimated likelihood approach

Proposed Innovation

Bayesian framework with Gibbs Sampler

- ▶ incorporate prior information
- ▶ fix/vary variance components
 - ▶ e.g., fix $\text{Var}(S(0))$ and assess sensitivity to different fixed values
- ▶ extend feasibility to settings with variable control-group response

Computational Strategy: Gibbs Sampler

Sample missing S

- ▶ draw missing $S(0)$ from distribution of $S(0)|X, Y, S(1)$
 - ▶ $S_i(0)|X_i, Y_i, S_i(1) \sim N(\mu_{i,0}, \sigma_{S(0)}^2)$
 - ▶ use information from observed $(X, S(0))$ in $Z = 0$ patients
- ▶ draw missing $S(1)$ from distribution of $S(1)|X, Y, S(0)$
 - ▶ $S_i(1)|X_i, Y_i, S_i(0) \sim N(m_{i,1}, \sigma_{S(1)}^2)$
 - ▶ use information from observed $(X, S(1))$ in $Z = 1$ patients
- ▶ use “complete” data to estimate parameters relevant to $CEP^{risk}(s_0, s_1)$

Outline

Introduction

- Potential Outcomes
- Posttreatment Variables
- Principal Stratification

Surrogate Endpoints

- Definition
- Evaluating Surrogates

Estimating Surrogate Value

- Causal Effect Predictiveness (CEP)
- Illustration**

Simulated Data

Underlying data (n=1000)

- ▶ Z induces a decrease in S
 - ▶ $\mu_{S(0)} = .41, \mu_{S(1)} = 0$
- ▶ X is very predictive of S
 - ▶ $\rho(X, S(0)) = .85, \rho(X, S(1)) = .85$
- ▶ $(S(0), S(1), X) \sim MVN(\mu, \Sigma)$

Simulated Data

Outcomes, Y

- ▶ $\Phi^{-1}(Pr(Y(0) = 1)) = \beta_{00} + \beta_{10}S(0) + \beta_{20}(S(1) - S(0))$
- ▶ $\Phi^{-1}(Pr(Y(1) = 1)) = \beta_{01} + \beta_{11}S(0) + \beta_{21}(S(1) - S(0))$
- ▶ $S(0)$ effect: $\beta_{10} = \beta_{11} = .5$
- ▶ no dissociative effect: $\beta_{00} = \beta_{01}$
- ▶ high associative effect: $\beta_{20} = 0, \beta_{21} = 2.5$

CEP Surface

$$CEP^{risk}(s_1, s_0) \equiv h(risk_{(1)}(s_0, s_1), risk_{(0)}(s_0, s_1))$$

▶ $h(x, y) = y - x$

$$CEP^{risk}(s_1, s_0) =$$

$$\Phi(\beta_{00} + \beta_{10}S(0) + \beta_{20}(s_1 - s_0)) - \Phi(\beta_{01} + \beta_{11}S(0) + \beta_{21}(s_1 - s_0))$$

Summary Descriptives

Intervention group ($Z = 1$)

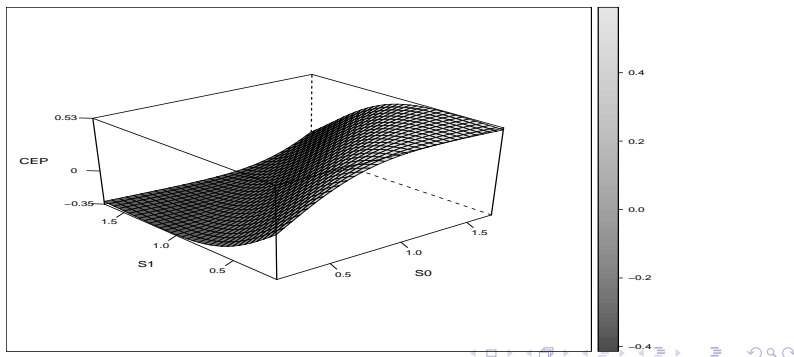
- ▶ $Pr(Y = 1|Z = 1) = 140/500 = .28$
- ▶ $\bar{S}(1) = 0$

Control group ($Z = 0$)

- ▶ $Pr(Y = 1|Z = 0) = 289/500 = .58$
- ▶ $\bar{S}(0) = .42$

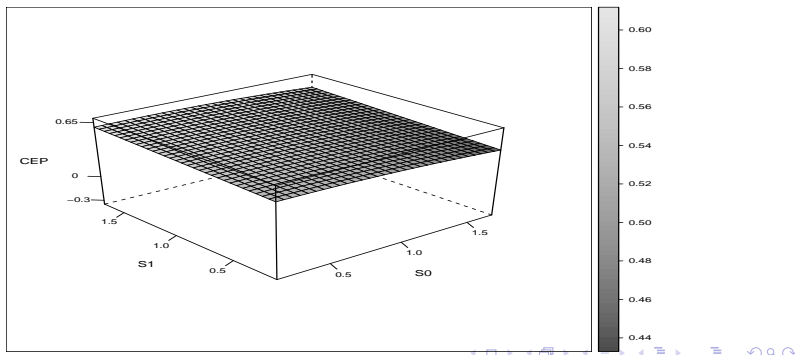
Simulated CEP Surface

Figure: CEP Surface with high surrogate value ($\beta_{00} = \beta_{01}, \beta_{21} = 2.5$)



Little Surrogate Value

Figure: CEP Surface with little surrogate value ($\beta_{00} > \beta_{01}, \beta_{21} = 0$)



Potential Questions of Interest

- ▶ Is it possible to collect/process biomarker information in far-flung field settings?
- ▶ How large a pilot study might be needed to anchor connections between surrogate and clinical endpoint?
- ▶ Would pilot-study findings from an English-language survey be applicable to non-English settings?
- ▶ Can a procedure be developed that is accessible to researchers who don't specialize in biostatistics?