Validly Estimating True Dose-Response When Only Treatment versus Control is Randomized: Principal Stratification for Causal Inference with Extended Partial Compliance

Hui Jin & Donald B. Rubin

Department of Statistics
Harvard University

December, 2006

Overview

Background: Efron and Feldman (1991) - EF:

- One of the earliest statistical articles to address non-compliance in randomized experiments.
- EF analyzed data from the Lipid Research Clinics Coronary Primary Prevention Trial (LRC-CPPT) to study the effectiveness of cholestyramine for lowering cholesterol levels.
- LRC-CPPT: Randomized treatment versus placebo, not dose.
- EF discussed inference for “dose-response” from non-randomized data.

Overview

Our work:

- Analyze the same data within the framework of Principal Stratification (Frangakis and Rubin, 2002).
- Explicate possible assumptions, including more flexible ones.
- Check EF's assumptions within our model.
- Formalize inference for dose-response.
- Our idea applies to any setting where dose is not randomized, e.g., amount of studying, hours of job-training.

LRC-CPPT Data

Specific features of LRC-CPPT:

- Placebo-controlled double blind randomized clinical trial to study the effectiveness of cholestyramine.
- 164 men were randomized to the treatment group and assigned the drug.
- 171 men were randomized to the control group and assigned placebo.
- For each patient, cholesterol levels were measured before and after taking the drug (or placebo).
- The outcome variable, Y, was the decrease in cholesterol level: the only variable available to EF or to us, besides treatment assigned and dose taken.

LRC-CPPT Data

Partial Compliance Complications:

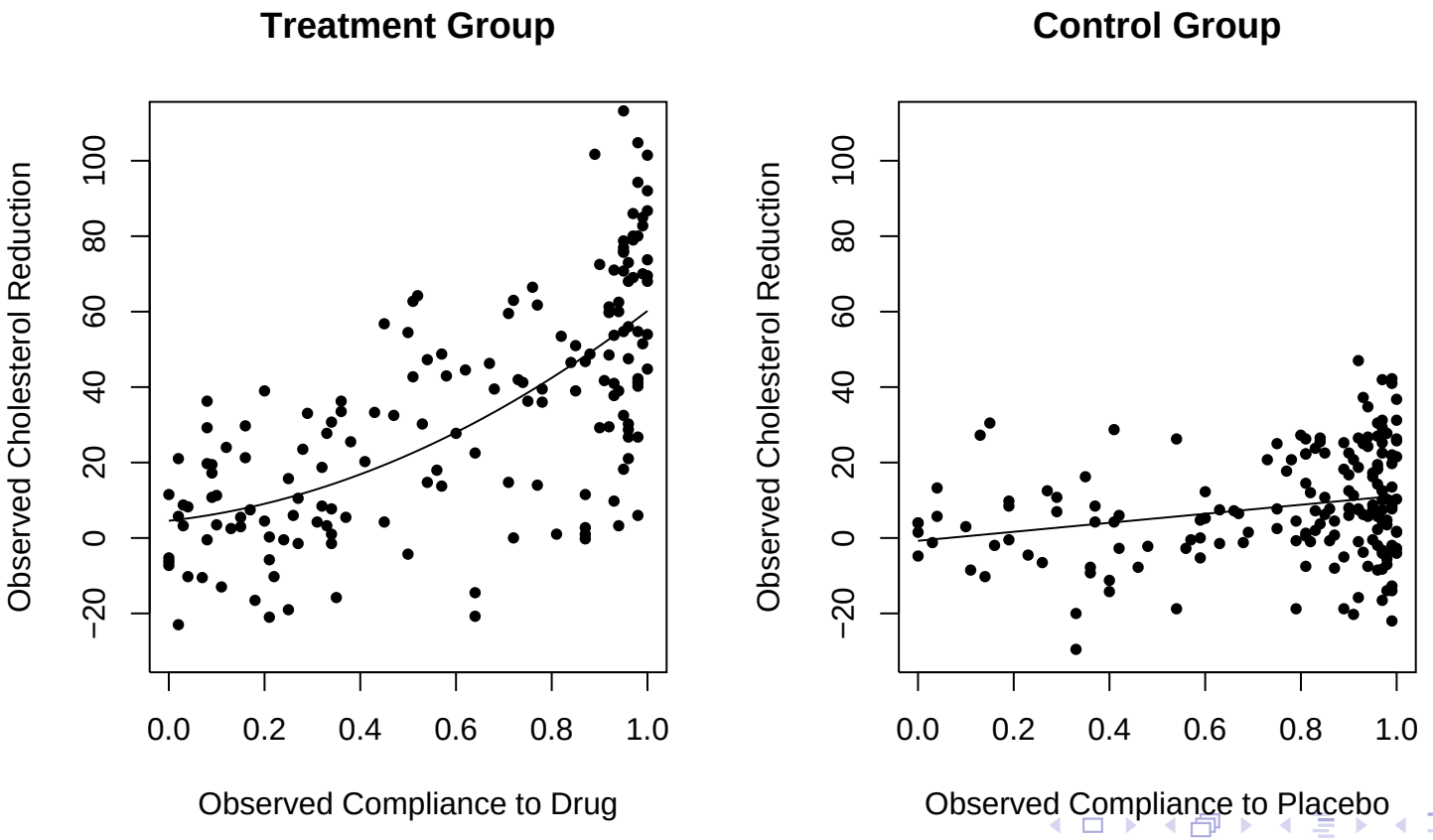
- Most patients in the treatment group only took a proportion of the assigned drug.
- Most patients in the control group only took a proportion of the assigned placebo.

Data Available:

- Z_i: treatment assignment
- $D_i(T)$ or $d_i(C)$: compliance to drug under treatment or compliance to placebo under control
- $Y_i(T)$ or $Y_i(C)$: outcome under treatment or outcome under control

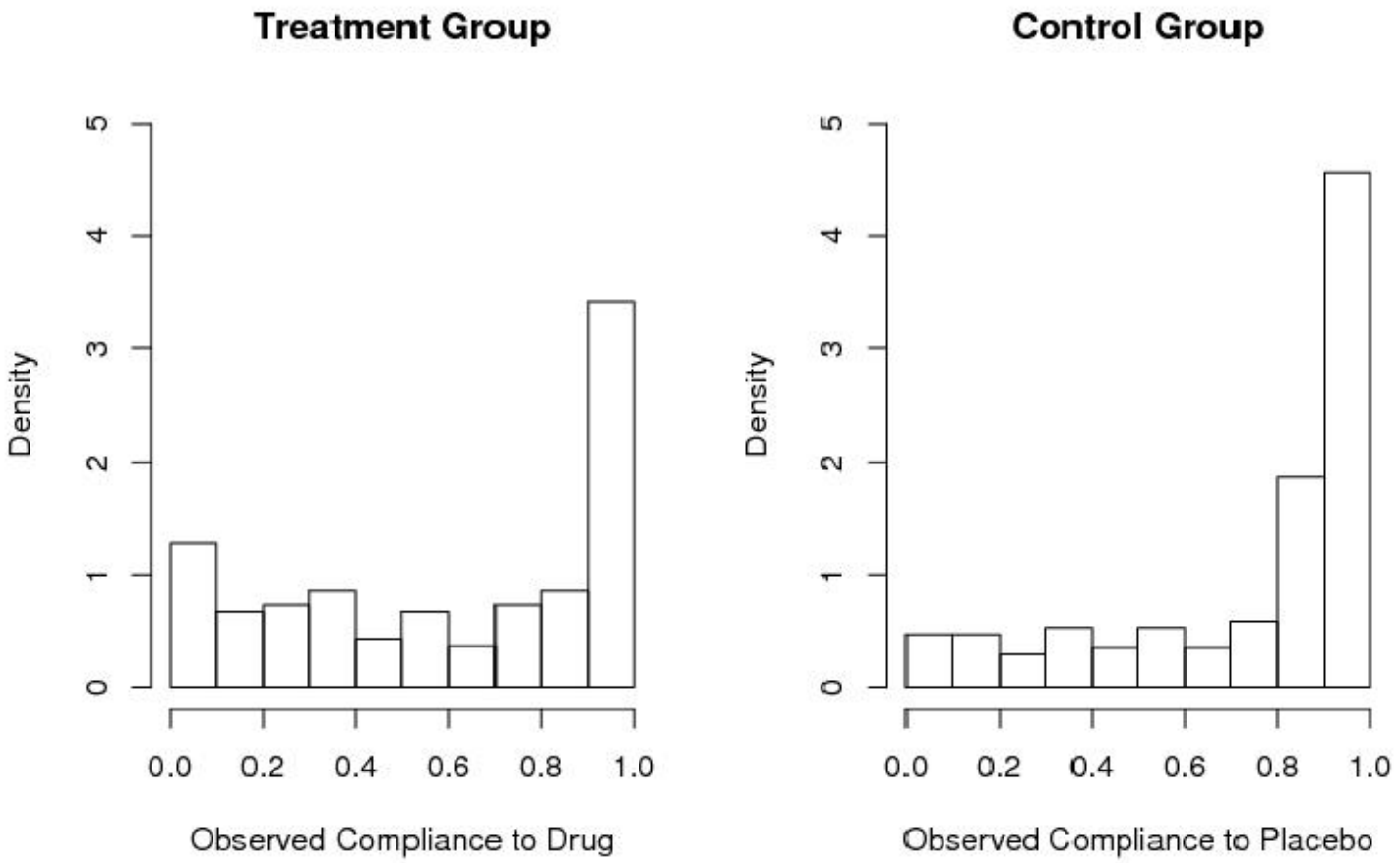
LRC-CPPT Data: Figure 1

Relationship Between Observed Cholesterol Reduction and Observed Compliance



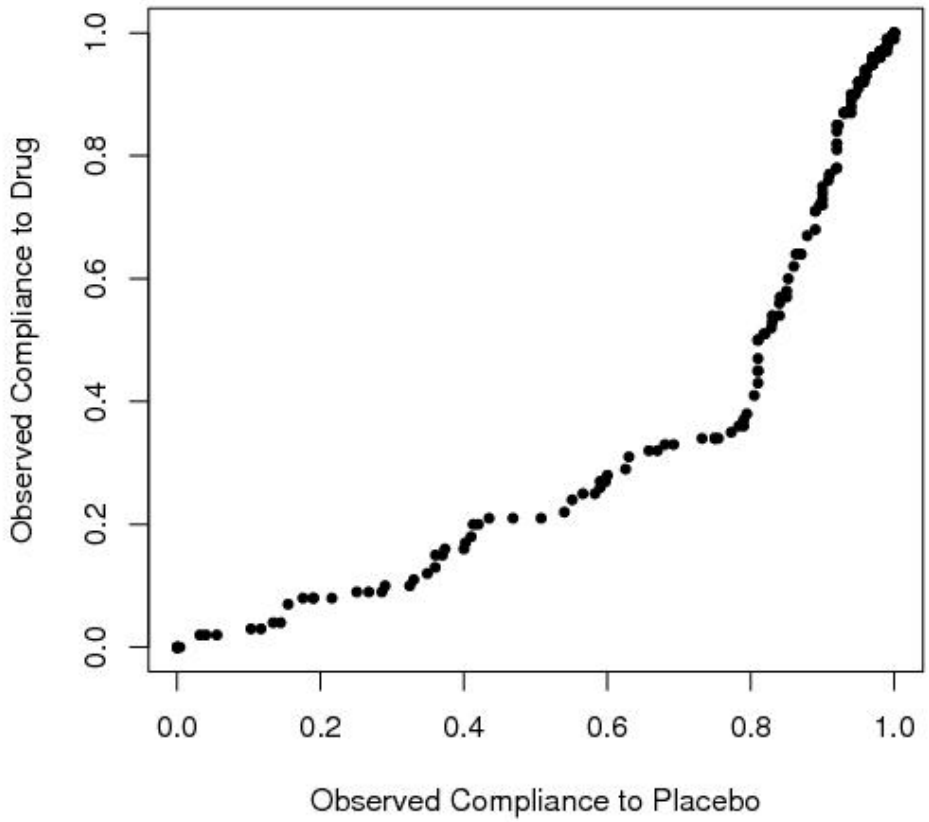
LRC-CPPT Data: Figure 2

Histograms of Observed Compliances



LRC-CPPT Data: Figure 3

Q-Q Plot of Observed Drug and Observed Placebo Compliances



Full Principal Stratification with Extended Compliance

i	X_i	Z_i	$D_i(T)$	$D_i(C)$	$d_i(T)$	$d_i(C)$	$Y_i(T)$	$Y_i(C)$
1	×	T	×	?	×	?	×	?
2	×	T	×	?	×	?	×	?
3	×	T	×	?	×	?	×	?
4	×	T	×	?	×	?	×	?
5	×	C	?	×	?	×	?	×
6	×	C	?	×	?	×	?	×
7	×	C	?	×	?	×	?	×
8	×	C	?	×	?	×	?	×

- **Individual Causal Effect**: $E_i = Y_i(T) - Y_i(C)$
- **Principal Stratum**: $S_i = [D_i(T), D_i(C), d_i(T), d_i(C)]$; "Full"⇒strata considered property of patients.
- **Principal Causal Effect**: $\overline{E}_s = AVE_{i \in S}[Y_i(T) - Y_i(C)]$. Average causal effect in principal stratum S.

Standard Assumptions

- **Stable Unit Treatment Value Assumption (SUTVA)**:
  One patient's treatment assignment will not affect other patients' potential outcomes;
  No hidden versions of treatment and no hidden versions of control.
- **Ignorable Treatment Assignment of T versus C**:
  True for randomized experiment.

These are accepted by both EF and us.

Assumptions at the Individual Level

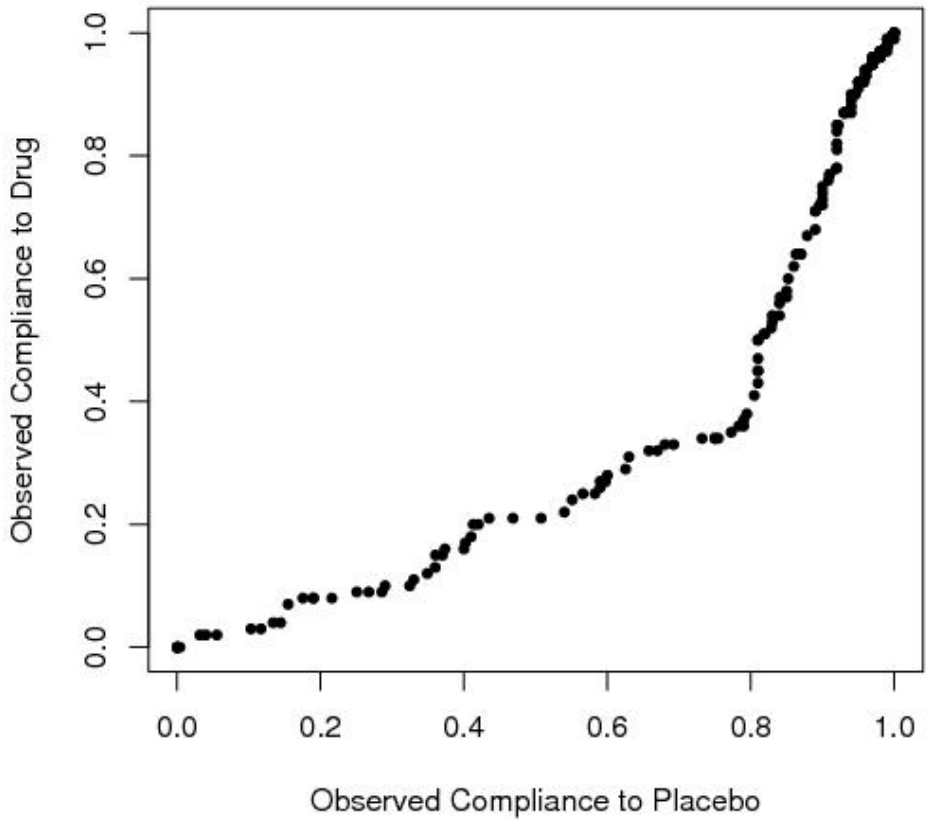
- **(1) Access Monotonicity**
  (1.A) General: $D_i(T) \geq D_i(C)$ and $d_i(C) \geq d_i(T)$
  (1.B) Strong: $D_i(C) = 0$ and $d_i(T) = 0$
- **(2) Side-Effect Monotonicity**
  (2.A) Negative: $D_i(T) \leq d_i(C)$
  (2.B) Positive: $D_i(T) \geq d_i(C)$
- **(3) Perfect Blind**: $D_i(T) = d_i(C)$
- **(4) Equipercentile Equating of Compliances**:
  $D_i(T) = F_D^{-1}\{F_d[d_i(C)]\}$
  Align percentiles of $D_i(T)$ and $d_i(C)$.

For LRC-CPPT:

- **EF assumed**: (1.B) and (4)$\Leftarrow$ true in expectation
- **We assume**: (1.B) and (2.A)$\Leftarrow$ weaker than (4)

EF's Assumption: Figure 3 Revisited

Q-Q Plot of Observed Drug and Observed Placebo Compliances



Full Principal Stratification for LRC-CPPT

i	Z_i	$D_i(T)$	$D_i(C)$	$d_i(T)$	$d_i(C)$	$Y_i(T)$	$Y_i(C)$
1	T	$\times$	0	0	?	$\times$	?
2	T	$\times$	0	0	?	$\times$	?
...	T	$\times$	0	0	?	$\times$	?
n_T	T	$\times$	0	0	?	$\times$	?
$n_T + 1$	C	?	0	0	$\times$	?	$\times$
$n_T + 2$	C	?	0	0	$\times$	?	$\times$
...	C	?	0	0	$\times$	?	$\times$
n	C	?	0	0	$\times$	?	$\times$

- **Principal Stratum**: $S_i = [D_i(T), 0, 0, d_i(C)] = [D_i, d_i]$ for notational simplicity.
- Recall, S_i is property of patients; modified later.
- **Principal Causal Effect**: $\overline{E}_s = AVE_{i \in S}[Y_i(T) - Y_i(C)]$. Average causal effect in principal stratum S.

Parametric Model

Parametric model:

- $d_i|\theta,\rho \sim Beta(\alpha_1,\alpha_2)$;
  $\frac{D_i}{d_i}|d_i,\theta,\rho \sim Beta(\alpha_3,\alpha_4)$.
- $Y_i(C)|D_i,d_i,\theta,\rho \sim N(\beta_0+\beta_1 D_i+\beta_2 d_i,\sigma_C^2)$.
- $Y_i(T)|Y_i(C),D_i,d_i,\theta \sim N[(\gamma_0+\gamma_1 D_i+\gamma_2 D_i^2+\gamma_3 d_i)+\rho\frac{\sigma_T}{\sigma_C}(Y_i(C)-\beta_0-\beta_1 D_i-\beta_2 d_i),(1-\rho^2)\sigma_T^2]$.
- Partial correlation ρ: sensitivity parameter.

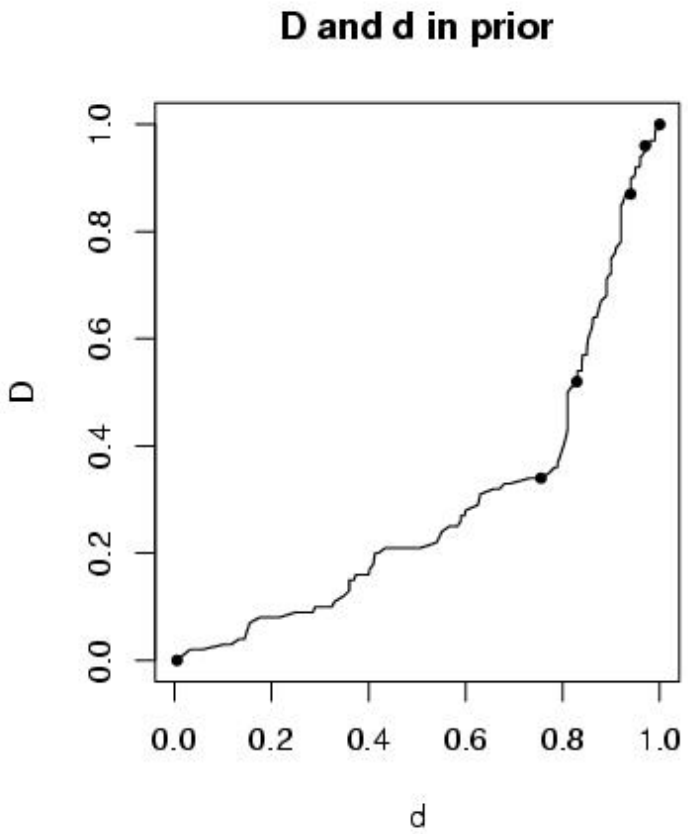
Therefore, $\theta=(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\beta_0,\beta_1,\beta_2,\gamma_0,\gamma_1,\gamma_2,\gamma_3,\sigma_C,\sigma_T)$.

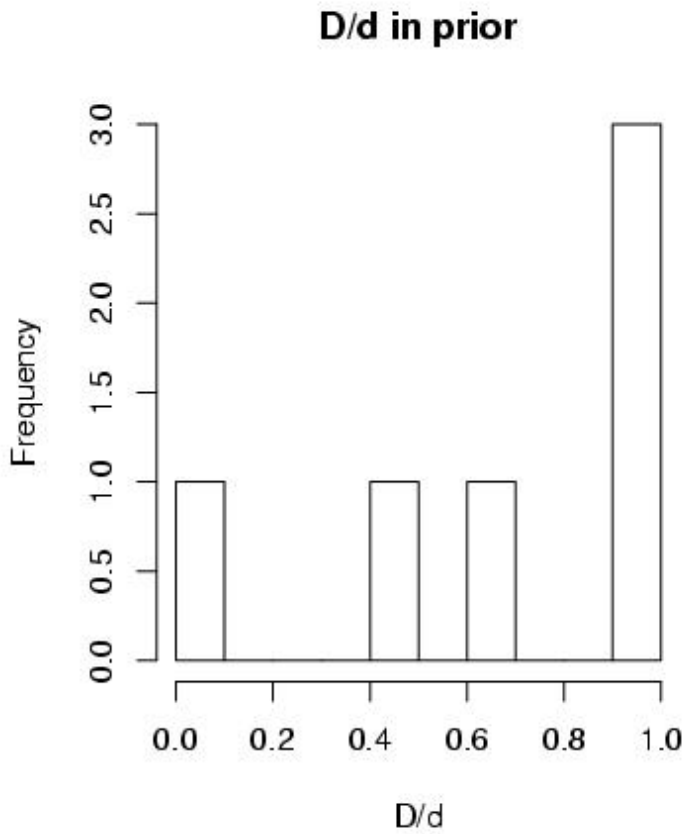
Prior distribution $\pi(\theta|\rho)$:

- $\pi(\alpha_1,\alpha_2,\alpha_3,\alpha_4|\rho)$: corresponds to adding 6 extra observations with complete (D,d) values.
- $\pi(\beta_0,\beta_1,\beta_2,\gamma_0,\gamma_1,\gamma_2,\gamma_3,\sigma_C,\sigma_T|\alpha_1,\alpha_2,\alpha_3,\alpha_4,\rho) \propto (\sigma_C\sigma_T)^{-2}$.

Parametric Model: Figure 4 - Prior Distribution

Prior Data Points for $\pi(\alpha_1, \alpha_2, \alpha_3, \alpha_4|\rho)$





Data Structure with Prior Data Points

i	Z_i	$D_i(T)$	$D_i(C)$	$d_i(T)$	$d_i(C)$	$Y_i(T)$	$Y_i(C)$
(1)	?	×	0	0	×	?	?
(2)	?	×	0	0	×	?	?
(...)	?	×	0	0	×	?	?
(6)	?	×	0	0	×	?	?
1	T	×	0	0	?	×	?
2	T	×	0	0	?	×	?
...	T	×	0	0	?	×	?
n_T	T	×	0	0	?	×	?
$n_T + 1$	C	?	0	0	×	?	×
$n_T + 2$	C	?	0	0	×	?	×
...	C	?	0	0	×	?	×
n	C	?	0	0	×	?	×

Computation: MCMC

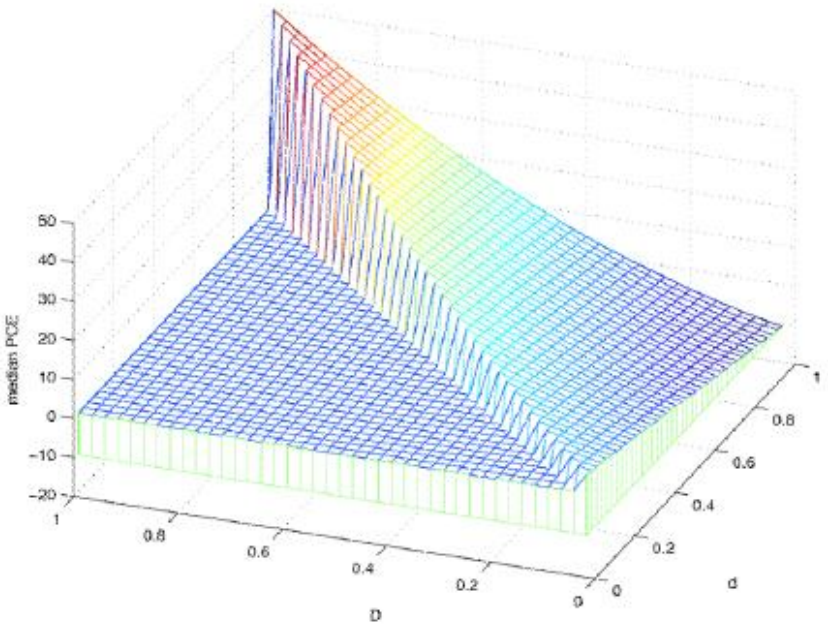
This is a missing data problem, and we can use MCMC to make Bayesian inference:

- **Parameters**: θ (fixed ρ)
- **Key missing data**: missing D_i or missing d_i
- In MCMC, given parameters θ, draw key missing data D_i or d_i; then given missing data, draw parameters; iterate until convergence.
- After convergence, we can simulate missing $Y_i(T)$ or $Y_i(C)$ and obtain a set of complete data.

Therefore, we can get the posterior distribution of every estimand: θ, principal causal effects, principal stratum of each patient,...

Figure 5 - Scientific Estimands

Principal Causal Effects:
Posterior Median and 95% Interval with $\rho = 0$



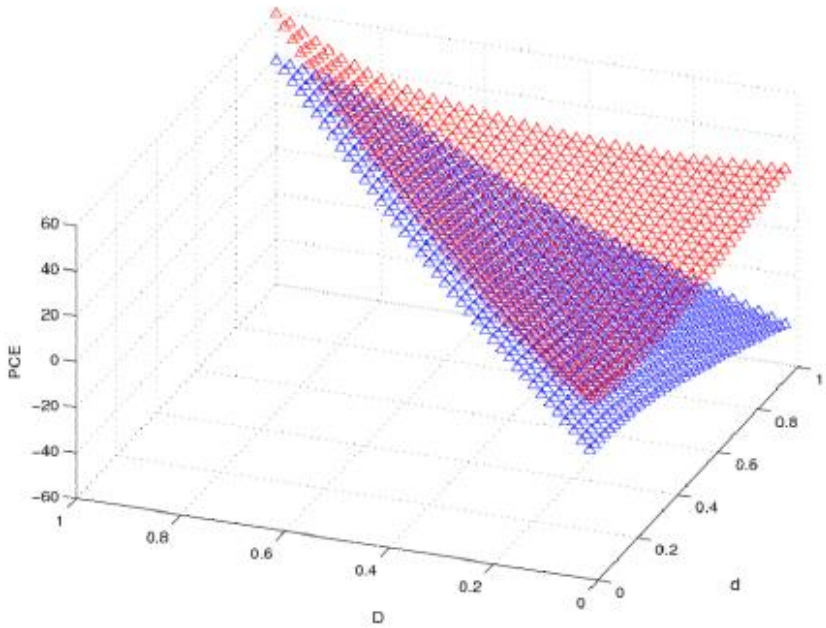


Figure 6 - Diagnostic Checks of EF's Assumption

Four Posterior Draws of Principal Strata for All the Patients with $\rho = 0$

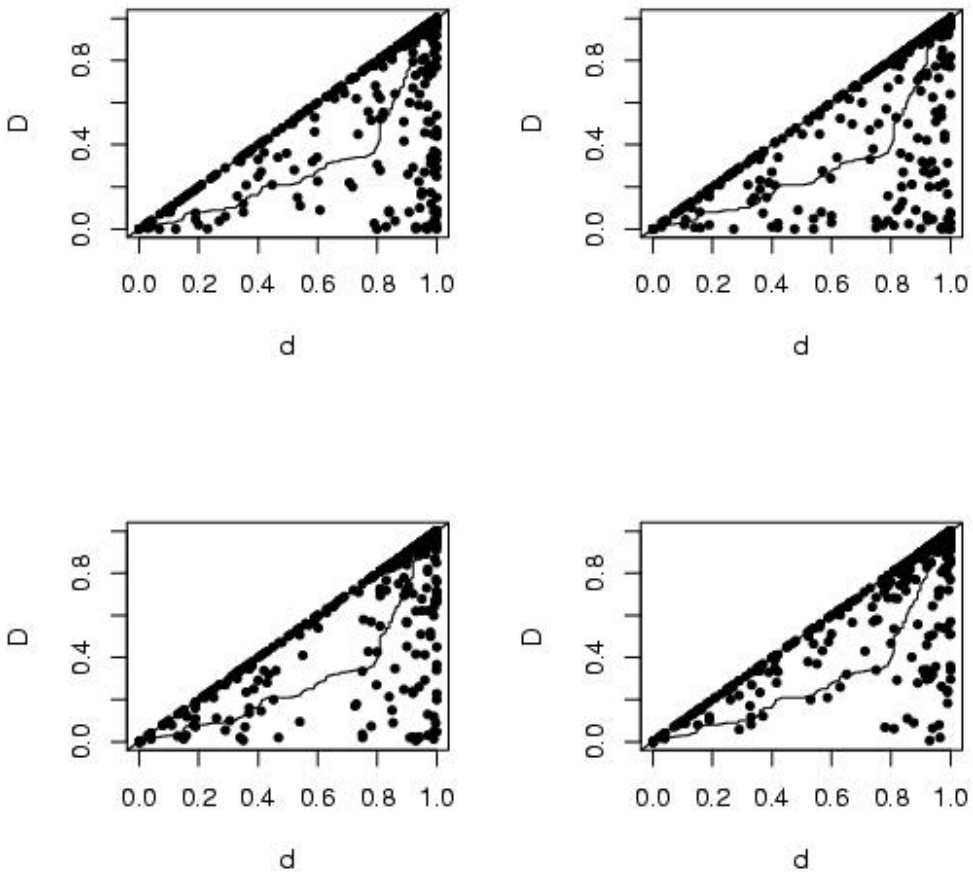
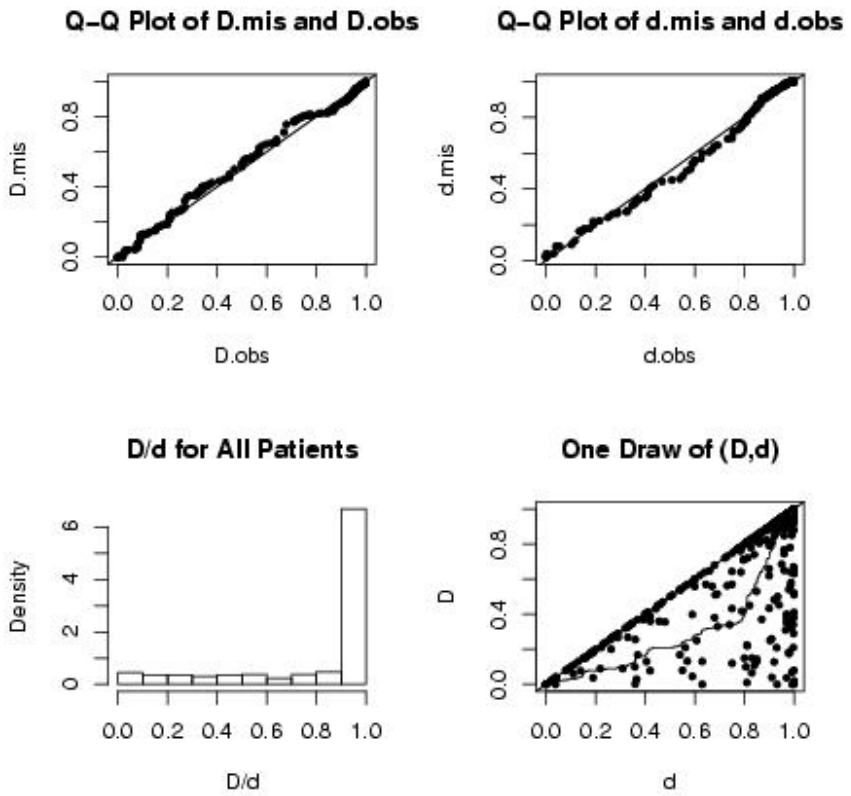


Figure 7 - Diagnostic Checks of Our Model

One Posterior Draw of D.mis and d.mis with $\rho = 0$



Sensitivity Analysis

Posterior Medians and Intervals of PCE for Different Values of ρ

(D, d)	(1, 1)	(0.68, 0.89)	(0, 1)	(0, 0)
$\rho = -0.2$	49 (39, 59)	24 (18, 30)	-10 (-40, 25)	4 (-6, 14)
$\rho = 0$	50 (39, 59)	24 (17, 30)	-13 (-42, 27)	5 (-6, 16)
$\rho = 0.2$	50 (39, 59)	23 (16, 29)	-11 (-47, 27)	5 (-7, 18)
$\rho = 0.4$	50 (40, 59)	23 (16, 29)	-6 (-43, 34)	6 (-7, 20)
$\rho = 0.6$	51 (39, 62)	22 (15, 30)	-10 (-43, 30)	7 (-8, 23)
$\rho = 0.8$	52 (38, 63)	22 (11, 33)	-8 (-62, 68)	6 (-11, 28)
$\rho = 0.9$	51 (37 ,66)	22 (6, 36)	-1 (-74 ,79)	9 (-25 ,41)

Understanding Principal Strata

Meaning of D_i and d_i:

- d_i: compliance to placebo indicates patient i's psychological compliance status.
- D_i: compliance to drug includes both patient i's psychological compliance status and his tolerance to negative side effects of the drug.
- d_i is more "fundamental" or "personal" than D_i.
- But D_i hints at possibility of estimating dose-response.
- Similar comments in EF.

Estimating the Dose-Response Relationship

Dose-Response within the Principal Stratification Framework:

- To estimate a dose-response relationship, we need a hypothetical experiment where different doses of drug are randomly assigned and strictly enforced (Also in EF).
- In the EF data, we need an additional assumption: for each cohort of patients with the same d, the assignment of D is stochastic and "latent ignorable" (Frangakis and Rubin 1999).
- With this additional assumption, we need a modified Principal Stratification framework, where D_i is no longer a stratum indicator.

Estimating the Dose-Response Relationship

Specific hypothetical experiment:

- Measure $d_i^* =$ baseline compliance for each patient.
- Randomly divide patients into Treatment and Control.
- In the treatment group, stochastically assign dose $Z_{Di} \leq d_i^*$ according to a certain "rule".
- In the control group, assign full placebo and measure d_i.
- We notice $d_i = d_i^*$ in the control group, then "lose" d_i^* in the control group and in the treatment group.
  $\Rightarrow$ Non-ignorable assignment of Z_{Di},
  ...but latent ignorable given d_i^*.
- Also, "forget" the rule for the assignment of Z_{Di}.

Estimating the Dose-Response Relationship

Modified Principal Stratification Framework for Dose-Response

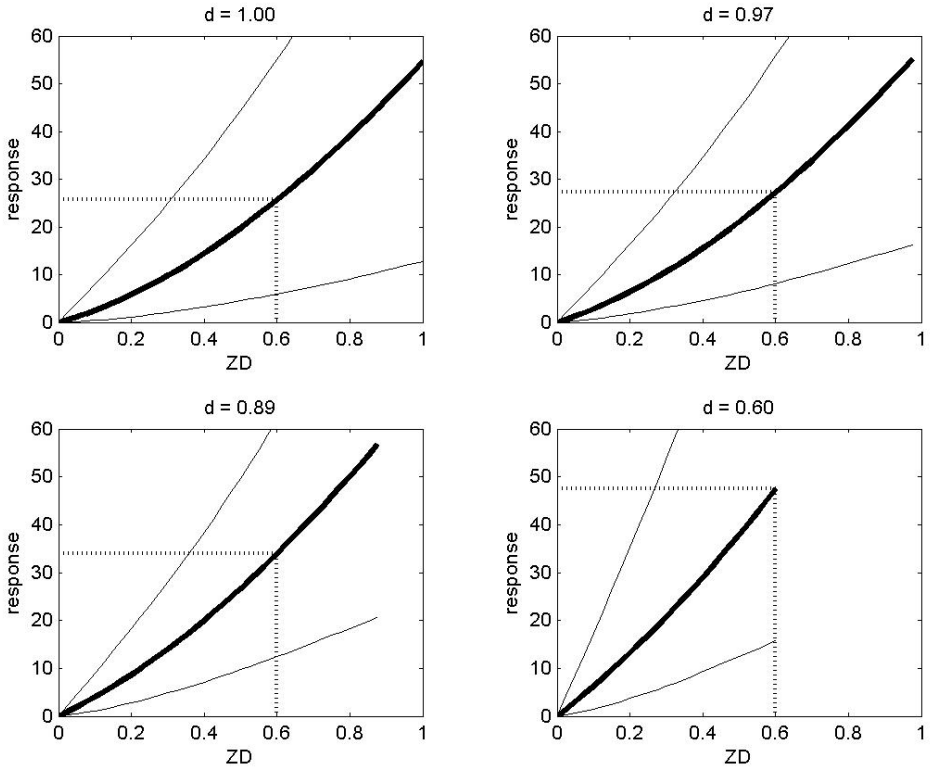
i	d_i^*	Z_i	Z_{Di}	$d_i(T)$	$d_i(C)$	$Y_i(T_0)$	...	$Y_i(T_D)$	...	$Y_i(T_1)$	$Y_i(C)$
1	?	T	T_0	0	?	$\star$	?	?	?	?	?
...	?	T	...	0	?	...	...	...	...	...	?
...	?	T	T_D	0	?	?	?	$\star$	?	?	?
...	?	T	...	0	?	...	...	...	...	...	?
n_T	?	T	T_1	0	?	?	?	?	?	$\star$	?
$n_T + 1$	?	C	?	0	$\star$	?	?	?	?	?	$\star$
...	?	C	?	0	$\star$	?	?	?	?	?	$\star$
...	?	C	?	0	$\star$	?	?	?	?	?	$\star$
n	?	C	?	0	$\star$	?	?	?	?	?	$\star$

"$\star$" represents observed data, "?" represents missing data.

Modified Parametric Model

- $d_i|\theta \sim Beta(\alpha_1, \alpha_2)$;
  $\frac{Z_{Di}}{d_i}|d_i, \theta \sim Beta(\alpha_3, \alpha_4)$.
  – Prior distribution remains "the same" for (Z_{Di}, d_i).
- $Y_i(C)|d_i, \theta \sim N(\beta_0 + \beta_2 d_i, \sigma^2_C)$.
  – No dependence on randomly assigned dose, Z_{Di}.
- $Y_i(Z_{Di})|Y_i(C), d_i, \theta \sim$
  $N[Y_i(C) + \gamma_1 Z_{Di} + \gamma_2 Z^2_{Di} + \gamma_3 Z_{Di} d_i, \sigma^2_{T.C}]$,
  where $\gamma_1 \geq 0, \gamma_2 \geq 0, \gamma_1 + \gamma_3 \geq 0$.
  – When $Z_{Di} = 0$, expectation of $Y_i(Z_{Di}) - Y_i(C) = 0$.
  – Dose-response is monotonely increasing.

Figure 8 - Dose-Response Results



Discussion of the Dose-Response Results

- Full Principal Stratification results are not causal for dose-response, but descriptive given principal strata.
- Dose-response results are causal under debatable assumptions.
- Is $Pr(Z_D|d,\{Y\}) = Pr(Z_D|d)$, "Nature's randomization" of dose Z_{Di} given d_i, plausible?
- Or do we need $Pr(Z_D|d,\{Y\},M) = Pr(Z_D|d,M)$, where M refers to medical side effects of the drug beyond d?
  $\Rightarrow$ Not latent ignorable given d, but latent ignorable given d and M.
- Sensitivity analysis to M? $\Rightarrow$ future work.

Main References

- Efron, B., Feldman, D. (1991), "Compliance as an Explanatory Variable in Clinical Trials," *Journal of the American Statistical Association* 86, 9-17.
- Frangakis, C.E. and Rubin, D.B. (2002), "Principal Stratification in Causal Inference," *Biometrics* 58, 20-29.
- Hirano, K., Imbens, G.W., Rubin, D.B. and Zhou X.-H. (2000), "Assessing the Effect of an Influenza Vaccine in an Encouragement Design," *Biostatistics* 1, 69-88.
- Frangakis, C.E. and Rubin, D.B. (1999), "Addressing Complications of Intention-to-Treat Analysis in the Combined Presence of All-or-None Treatment - Noncompliance and Subsequent Missing Outcomes," *Biometrika* 86, 365-379.