Simple Noncompliance, Instrumental Variables, and Bayesian Generalizations

- Template for other observational studies involves more complex randomized experiment
- Illustrate with completely randomized experiment with noncompliance with assigned treatment
- Return later to combined analysis with observational study design

Sommer and Zeger Vitamin A Data

Row	True Compliance Type	Treatment Assignment	Treatment Received	Y_{obs}	Number of Children
1	?	0	0	0	11514
2	?	0	0	1	74
3	N	1	0	0	2385
4	N	1	0	1	34
5	C	1	1	0	9663
6	C	1	1	1	12
					23682

Reference: Sommer and Zeger (1991). On Estimating Efficacy from Clinical Trials. Statistics in Medicine.

Results of Three Standard MoM Analyses

Method	Estimate	Calculation	Row Comparison
ITT	-0.0026		3, 4, 5, & 6 vs. 1 & 2
As-treated	-0.0065		5 & 6 vs. 1, 2, 3, &4
Per protocol	-0.0052		5 & 6 vs. 1 & 2

Reference: Sommer and Zeger (1991). On Estimating Efficacy from Clinical Trials. Statistics in Medicine.

MoM CACE Analysis

ACE = $p_N \cdot$ NACE + $p_C \cdot$ CACE

-0.0025 = 0.2 · NACE + 0.8 · CACE

-0.0025 = 0.8 · CACE ➔ CACE = -0.0025/0.8 = -0.0031

Bayesian Analysis of Sommer & Zeger Data

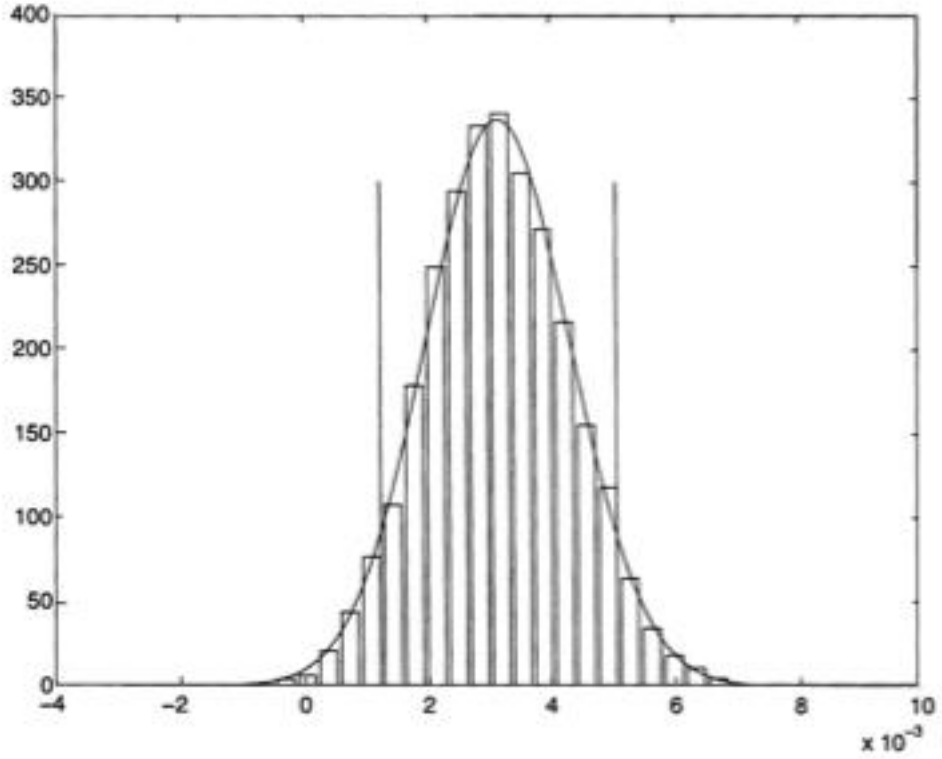


FIG. 3. *Histogram of CACE with exclusion restriction (data from Table 3).*

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Bayesian Analysis of Sommer & Zeger Data

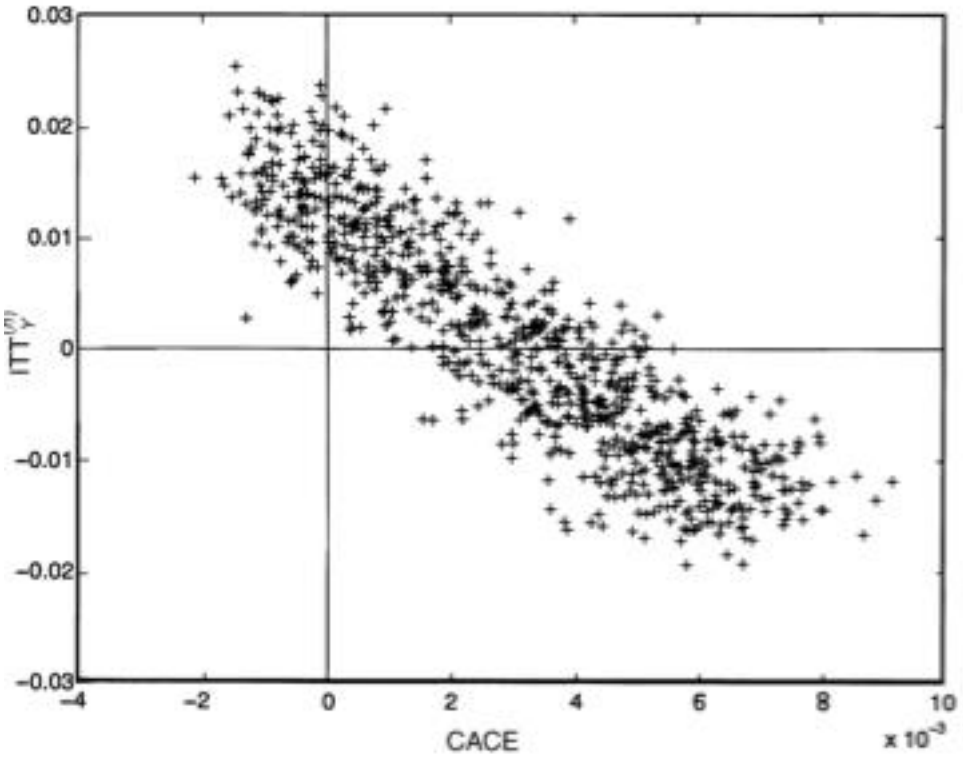


FIG. 4. *Joint posterior distribution of CACE and* $\mathrm{ITT}_Y^{(c)}$ *(data from Table 3).*

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Hypothetical Example Illustrating Frequentist Superiority of Bayes over IVE (MoM) and MLE, Population Parameters with Exclusion Restrictions and Monotonicity

T	$P(C_i = t \mid \pi)$	$D_i(0)$	$D_i(1)$	$Y_i \mid C_i = t, Z_i = 0, \pi$	$Y_i \mid C_i = t, Z_i = 0, \pi$
c	0.25	0	1	$N(0.1, 0.16)$	$N(0.9, 0.49)$
n	0.45	0	0	$N(1.0, 0.25)$	$N(1.0, 0.25)$
a	0.30	1	1	$N(0.0, 0.36)$	$N(0.0, 0.36)$

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

Hypothetical Example Illustrating Frequentist Superiority of Bayes over IVE (MoM) and MLE, One Sample

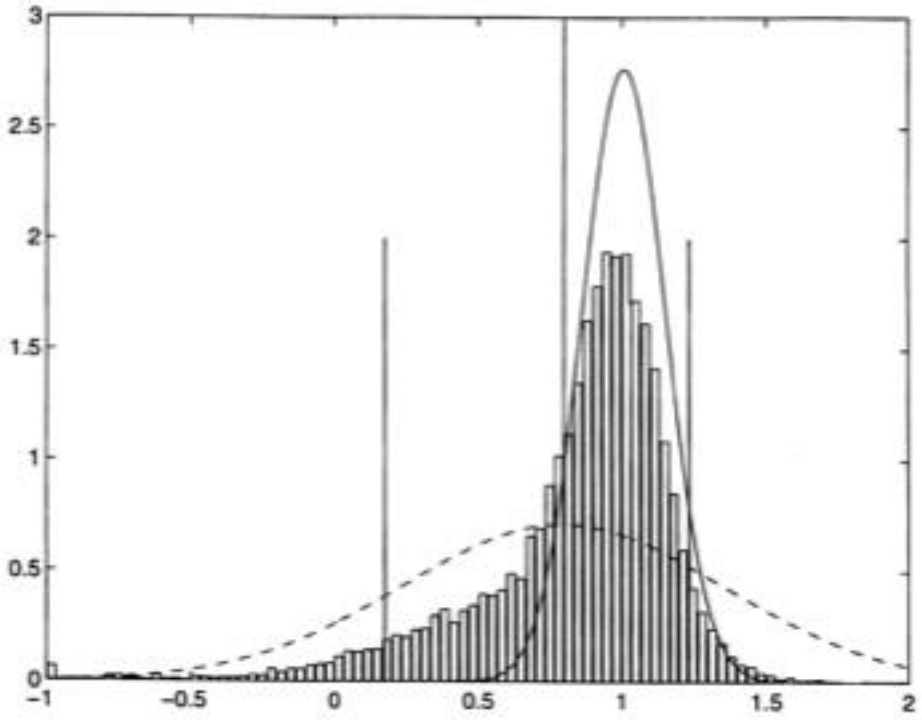


FIG. 5. *Estimates of the posterior distribution of CACE under exclusion restriction and monotonicity condition (data analyzed in Table 6): histogram is based on simulation, solid line is normal approximation based on mle, dashed line is normal approximation based on* $\widehat{\text{IVE}}$.

Imbens G.W. and Rubin D.B. (1997) Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. Annals of Statistics 25(1):305-327.

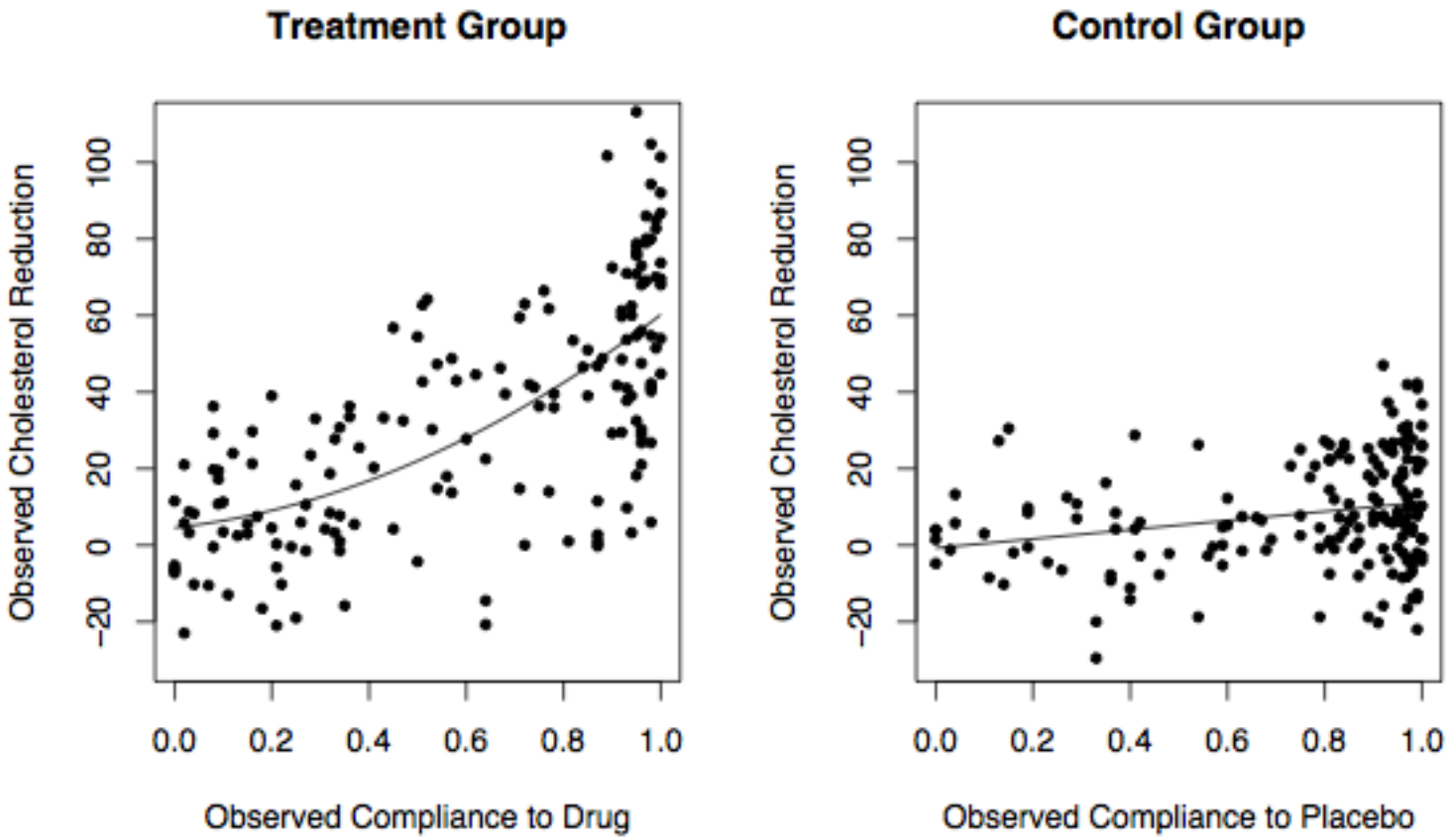
The EF Data

- 164 men were randomized to the treatment group and assigned the drug, $Z_i = T$
- 171 men were randomized to the control group and assigned placebo, $Z_i = C$
- For each patient, cholesterol levels were measured before and after taking the drug or placebo
- The outcome variable, $Y_i(T)$ or $Y_i(C)$, was the decrease in cholesterol level: the only variable used by EF or JR, besides treatment assigned and dose taken

Complications with the EF Data

- Partial and Extended Noncompliance
  - Most patients in the treatment group took only a proportion of the assigned drug: $D_i(T) \in [0,1]$
  - Most patients in the control group took only a proportion of the assigned placebo: $d_i(C) \in [0,1]$
  - By design, $D_i(C) = 0$ and $d_i(T) = 0$
- Thanks to Brad Efron for sharing data

Relationship Between Observed Cholesterol Reduction and Observed Compliance



Figures from Efron and Feldman, 1991

How to estimate dose-response?

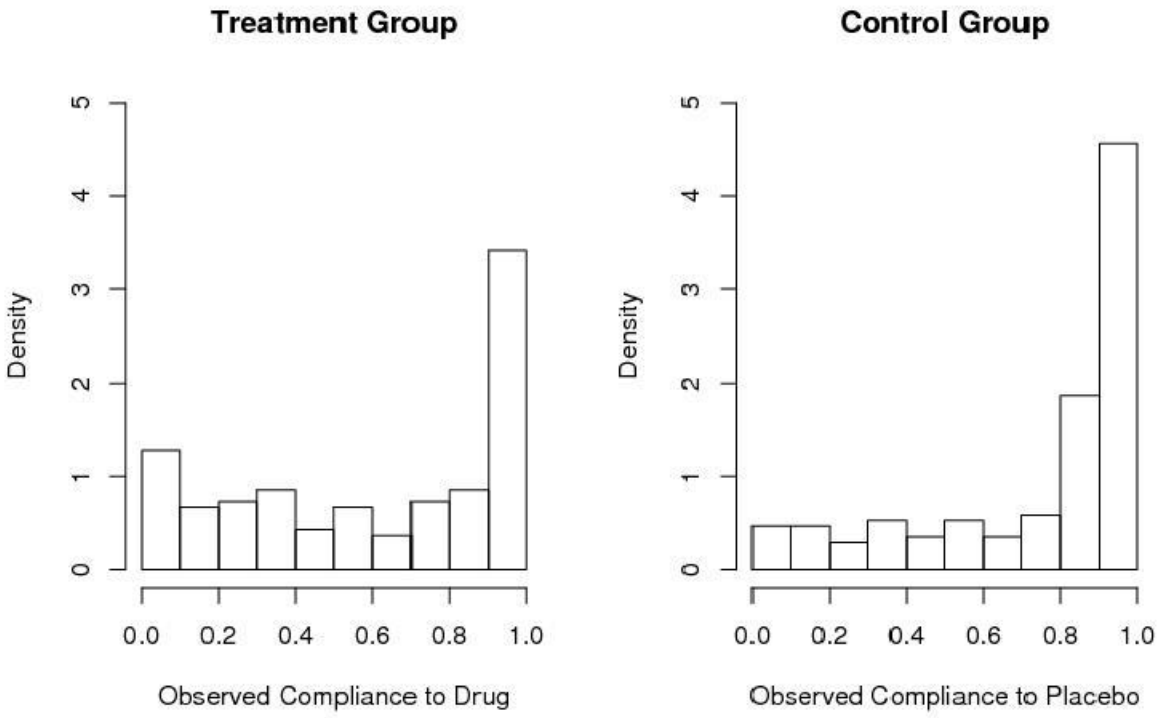
- Observed “dose-response” in both arms
- Somehow, “subtract” Y(C) versus d(C) plot from Y(T) versus D(T) plot
- EF attempted this, but written discussion (including by DBR) of article indicated debatable success
- Objective in JR was to do this “subtraction” correctly under explicit assumptions
- Here, highlight principal stratification, hypothetical experiment for dose-response, and Bayesian approach to analysis

Standard Assumptions

- Stable Unit Treatment Value Assumption (SUTVA):
  - One patient's treatment assignment does not affect other patients' potential outcomes;
  - For each patient, 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 JR.

Histograms of Observed Compliance



Figures from Jin and Rubin, 2008

Q-Q Plot of Observed Drug and Observed Placebo Compliance

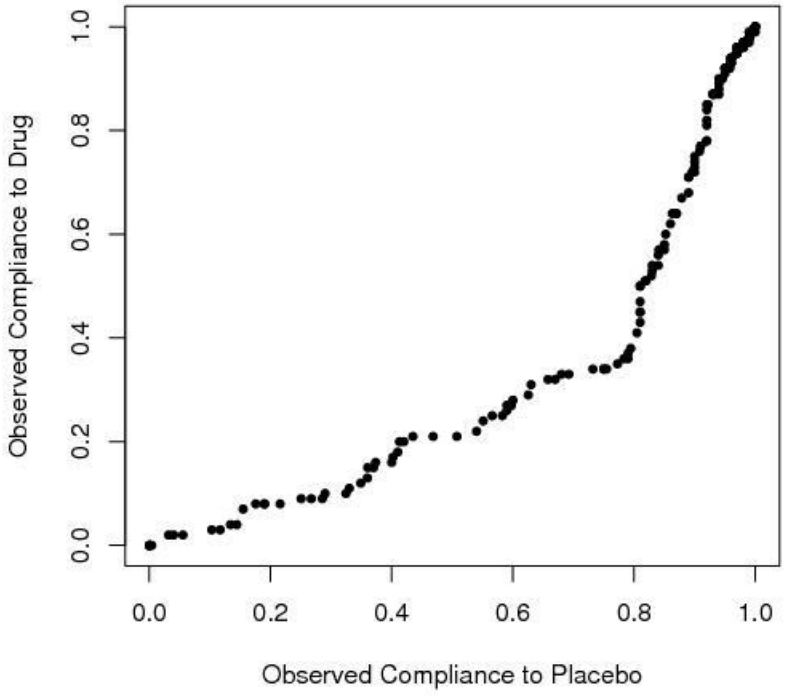


Figure from Jin and Rubin, 2008

Possible Assumptions at the Individual Level

- Perfect Blind: $D_i(T) = d_i(C)$; obviously wrong
- Equipercentile Equatable Compliances
  - Align percentiles of $D_i(T)$ and $d_i(C)$, as in Q-Q
  - So both known for all men: $D_i(T) = F(d_i(C))$
  - EF assume this, which is true in expectation
- Side-Effect Monotonicity
  - Negative: $D_i(T) \leq d_i(C)$
  - Positive: $D_i(T) \geq d_i(C)$
  - JR assume negative side effects; plausible

Meaning of d_i and D_i

- d_i: compliance to placebo indicates patient i's "psychological" compliance status, a covariate that is missing for men assigned drug
- D_i: compliance to drug reflects both patients i's psychological compliance status and his tolerance to negative side effects of the drug, etc.
- But D_i hints at possibility of estimating dose-response
- Similar comments in EF, but JR allow $D_i(T) \neq F(d_i(C))$

Estimating a Dose-Response Relationship within the "Rubin Causal Model" (Holland, 1986)

- To estimate a dose-response relationship
  - Need a hypothetical experiment where different doses of drug are randomly assigned and enforced
- Principal stratification framework (Frangakis and Rubin, 2002) – vast generalization of IVE
  - The intermediate outcome $d_i(C)$ is unaffected by treatment assignment
  - Therefore is a partially observed covariate
- For each stratum of patients with the same $d_i(C)$, the assignment of dose is stochastic and "latent ignorable" (Frangakis and Rubin, 1999)

Specific Hypothetical Experiment

- Measure d_i^*= baseline compliance for each patient when assigned full placebo dose
- Randomly divide patients into Treatment and Control
- In treatment group, stochastically assign dose $Z_{D_i} \leq d_i^*$ according to a Beta random variable
- In control group, assign full placebo and measure d_i
- We notice $d_i(C) = d_i^*$ in the control group, then "toss" d_i^* in the control group and in the treatment group
- Thus, nonignorable assignment of Z_{D_i}, but latent ignorable given d_i^*
- Also, "forget" the rule for the assignment of Z_{D_i}

Principal Stratification Framework for Dose-Response with $d_i(C)$ Defining Strata and $Z_{D_i}(T)$ Defining Dose

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	$\star$	?	?	?	?	?	$\star$

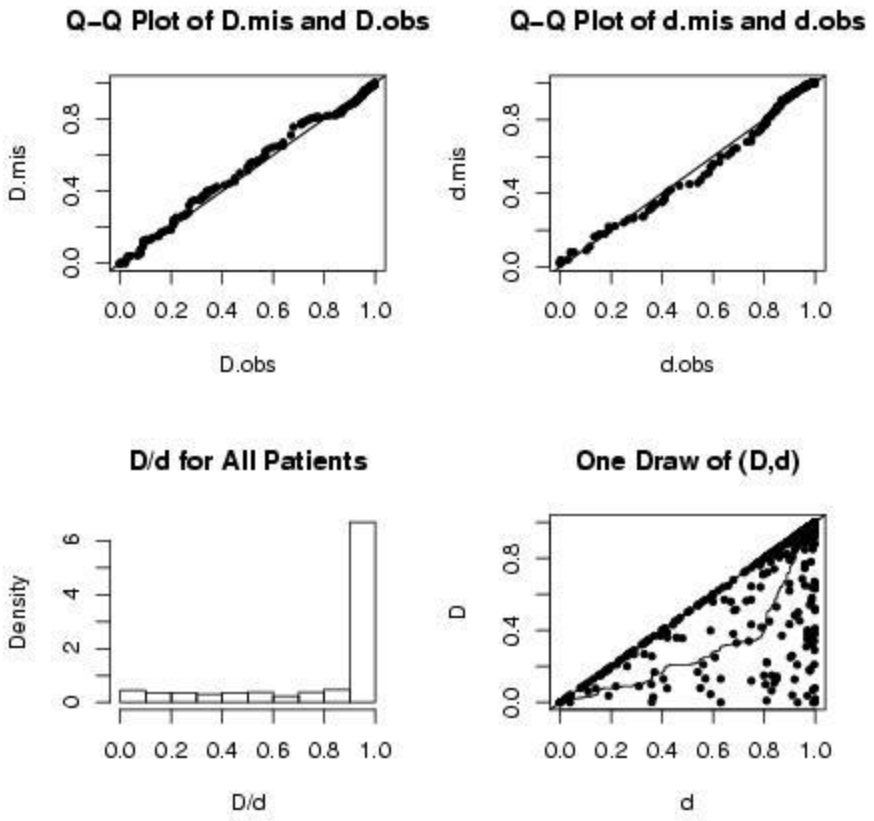
"$\star$" represents observed data, "?" represents missing data.

JR's Computation

- Missing data problem, which is addressed using MCMC to draw Bayesian inferences
  - Parameters are θ
  - Key missing data are d_i for those assigned treatment and Z_{D_i} for those assigned control
  - Given θ, draw key missing data; given key missing data, draw θ; iterate until approximate convergence
  - Vast number of such draws approximates posterior distribution of dose-response as a function of principal strata defined by $d_i(C)$

Diagnostic Checks for JR's Model

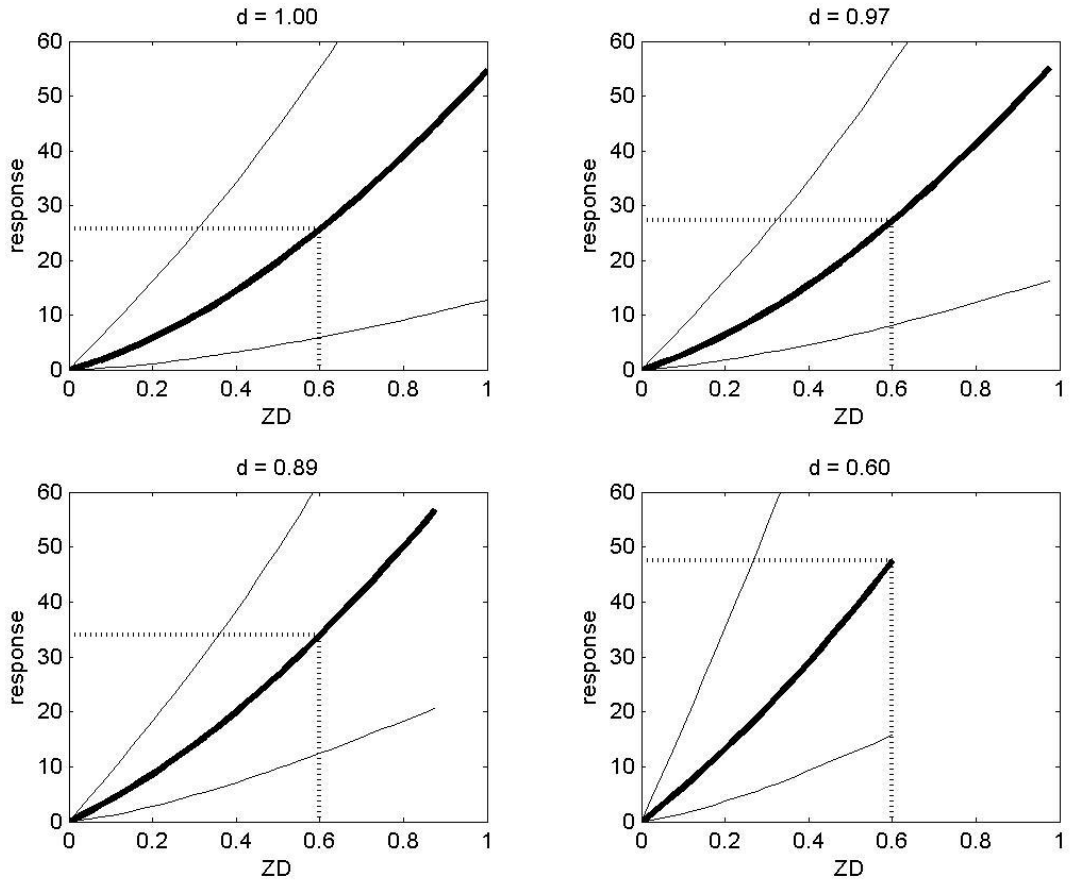
One Posterior Draw of Key Missing Data



Figures from Jin and Rubin, 2008

Dose-Response Results for Principal Strata

Maximum d, 75th d, median d, 25th d



Figures from Jin and Rubin, 2008

Discussion of the Dose Response Conclusions

- Under EF's assumptions, dose-response at each $d_i(C)$ is a point because $D_i(T) = F(d_i(C))$
- JR's dose-response results are causal under debatable assumption
  - Is "Nature's randomization" of dose given placebo compliance (i.e., the crucial latent ignorability assumption) plausible?
  - Or do we need to condition further on background medical characteristics related to possible side effects of the drug? Such sensitivity analysis is future work
- Framework is general and revealing, and much was anticipated earlier