Complier Average Causal Effect (CACE)

Booil Jo
Stanford University

Methodological Advancement Meeting
Innovative Directions in Estimating Impact

Office of Planning, Research & Evaluation
Administration for Children and Families
Washington DC, September 6-7, 2012

Randomized controlled trial (RCT)

- Gold Standard – strong causal inference (evidence) with the least bias based on random assignment of treatments and blinding.
- Typically compares an experimental treatment to a standard treatment or to a placebo – there are at least two groups of subjects.
- Random assignment results in unbiased distribution of confounding variables (e.g., gender, age, education). Blinding results in unbiased distribution of placebo effect. Treatment+placebo – placebo = treatment effect

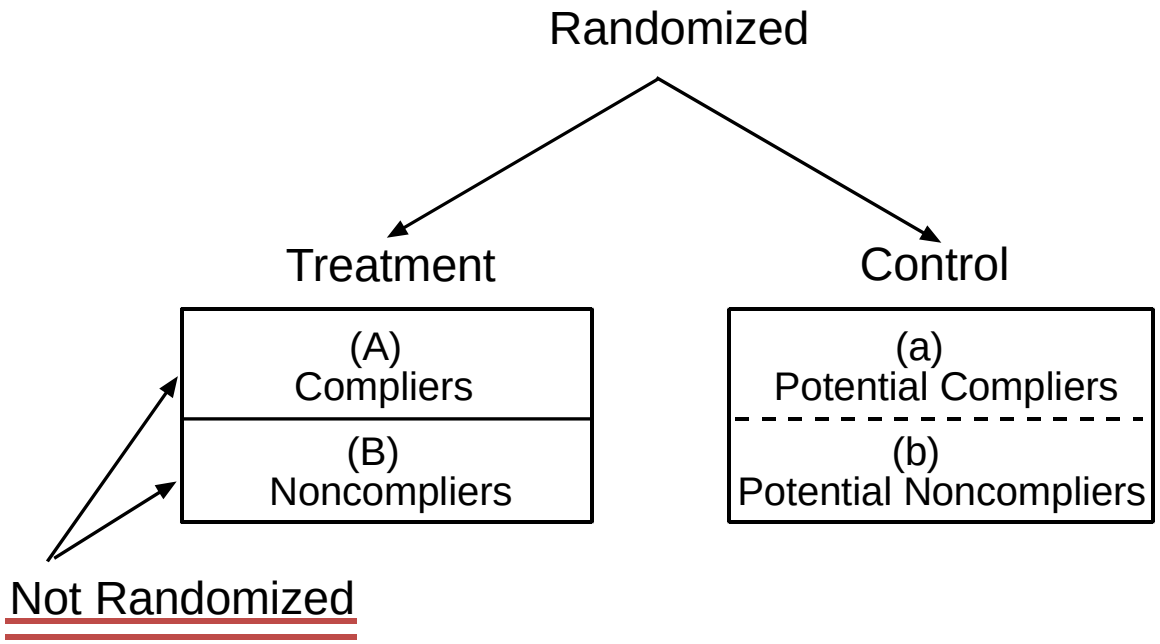
Possible Complications

- There could be many expected/unexpected complications that may undermine the validity of intervention effect estimates *despite randomization*:
  - Treatment noncompliance
  - Missing data (attrition, dropout, nonresponse)
  - Interference (within and across treatment groups)
  - Our desire to understand the intervention mechanism (mediation)
  - More…
- Our inference may vary substantially depending on how and how well these complications are handled.

Treatment Noncompliance

- Even with successful randomization and double blinding, estimating causal effects of treatments can be complicated by noncompliance.
- Noncompliance causes analysis complications because compliance status is often incompletely observed – can be seen as a missing data problem.
- Different conclusions may be reached about the effect of the same treatment depending on how noncompliance is dealt with in the analyses.

Treatment Noncompliance



- ITT Analysis: (A+B) vs. (a+b)
- As-Treated Analysis: (A) vs. (B+a+b)
- Per-Protocol Analysis: (A) vs. (a+b)

Example: Johns Hopkins School Intervention Study

- A randomized study designed to improve academic achievement and to reduce early behavioral problems (Ialongo et al., 1999).
- Among students assigned to the control condition, treatment receipt status could not be observed.
- In the Family-School Partnership intervention condition, parents were asked to implement 66 take-home activities.
- When the receipt of intervention treatment is defined as completing about 2/3 of activities (45 out of 66), about 47% of kids in the intervention condition properly received the intervention treatment.

Intention to Treat (ITT) Principle

- Trust the balance (comparability) based on random assignment
- Compare groups as randomized (as intended)
  - Compare means, proportions, trajectories, etc.
- Gold standard for analyzing data from RCTs
- ITT analysis presents a fair comparison and provides unbiased estimates of overall effectiveness of the treatment (in the absence of other complications).

Potential Outcomes Approach

(e.g., Neyman, 1923; Rubin, 1978, 1980…)

- The effect of treatment is defined based on an idealized (potential) situation, in which each individual's outcome is simultaneously observed under all compared conditions.
- $Y_i(1)$: potential outcome for individual *i* when assigned to the treatment condition ($Z = 1$)
- $Y_i(0)$: potential outcome for individual *i* when assigned to the control condition ($Z = 0$)
- $Y_i(1) - Y_i(0)$: the effect (ITT) of treatment assignment for individual *i*

Causal Interpretation of ITT Effect

- The individual level treatment effect, $Y_i(1) - Y_i(0)$, is interpreted as causal given that the only cause of the difference is the treatment assignment status.
- The individual level treatment effect $Y_i(1) - Y_i(0)$ generally cannot be identified
- The causal effect of treatment assignment can be defined at the average (population) level: $\mu_1 - \mu_0$
- The average causal effect $\mu_1 - \mu_0$ can be identified under certain conditions

1. Ignorability of Treatment Assignment

(Holland, 1986; Rubin, 1974, 1978, 1980)

- Key assumption that opens up possibilities of causal inference at the average level based on observed data.
- Treatment assignment is independent of the potential outcomes (given the observed covariates) – automatically satisfied in randomized experiments
- Individuals assigned to different treatment conditions have similar characteristics (comparable).

2. Stable Unit Treatment Value (SUTVA)

(Rubin, 1978, 1980, 1990)

- Another critical assumption that makes identification of causal treatment effects possible.
  - SUTVA I: potential outcomes for each person are unaffected by the treatment assignment of other individuals – “no interference”
  - SUTVA II: there is only one version of each treatment (who delivers the treatment does not affect the potential outcomes)

Interest in Treatment Efficacy

- As-treated and per-protocol analyses do not present a fair comparison and the resulting estimate cannot be interpreted as the causal effect of treatment.
- When ITT analysis is used, treatment efficacy tends to be understated by including individuals who did not take the treatment (considered conservative).
- Sometimes our interest is in estimating the effect of treatment when it is actually received.

Complier Average Causal Effect (CACE)

- A widely used alternative method of estimating treatment efficacy.
- Bloom (1984): instrumental variable (IV) approach
- Angrist, Imbens & Rubin (1996): IV+ potential outcomes
- Frangakis & Rubin (2002): generalize the AIR approach

A Widely Known Setting

- Angrist, Imbens, & Rubin, 1996.
- Random assignment to either to the treatment (Z=1) or to the control condition (Z=0).
- Individuals either receive (M=1) or do not receive (M=0) the treatment.
- Let $M_i(1)$ denote the potential treatment receipt status for individual i when assigned to the treatment condition (Z=1) and $M_i(0)$ when assigned to the control condition (Z=0).

Four Compliance Types (AIR 1996)

- Individuals are cross-classified into 4 latent compliance types (*Ci*) based on their *potential* treatment receipt behavior *M* under *both* treatment conditions (Z=0 or 1)

$$C_i = \begin{cases} n \text{ (never-taker)} & \text{if } M_i(1) = 0, \text{ and } M_i(0) = 0, \\ d \text{ (defier)} & \text{if } M_i(1) = 0, \text{ and } M_i(0) = 1, \\ c \text{ (complier)} & \text{if } M_i(1) = 1, \text{ and } M_i(0) = 0, \\ a \text{ (always-taker)} & \text{if } M_i(1) = 1, \text{ and } M_i(0) = 1. \end{cases}$$

- The causal treatment effect for compliers is widely known as complier average causal effect (CACE).
- Principal stratification (Frangakis & Rubin, 2002) can be thought of as generalization of this strategy.

Principal Stratification

(Frangakis & Rubin, 2002)

- It aims for causal inference accounting for subpopulation heterogeneity in terms of posttreatment intermediate outcomes (mediators).
- It means stratifying individuals based on potential values of a mediator under all treatment conditions.
- As a result, the principal stratum membership is independent of treatment assignment just like pretreatment baseline covariates.
- In each principal stratum, outcome of interest can be compared across treatment conditions – principal effect. Any principal effect is a causal effect.

Identification of CACE

- We cannot identify individuals' compliance type based on observed data
- To make up for this missing information and to make causal inference, we need to introduce external assumptions – these are also unverifiable!
- How do we come up with good assumptions?
  - Science
  - Experts opinion
  - Go with the flow (use assumptions that tend to be widely accepted)

Widely Used Identifying Assumptions
(Angrist, Imbens, & Rubin, 1996)

- To identify CACE, two identifying assumptions are employed in addition to ignorable treatment assignment and SUTVA
  - *Exclusion restriction*
  - *Monotonicity*

3. Monotonicity

- It excludes the possibility of having defiers.
- It is often a critical assumption that supports the identifiability of causal inference models (e.g., AIR).
- Considered a benign assumption in RCTs where we can control treatment receipt by study design. E.g., by disallowing individuals from receiving treatments other than those they are assigned to receive.
- It is harder to know if monotonicity is reasonable when the intermediate variable is not treatment receipt (e.g., quitting smoking).

4. Exclusion Restriction

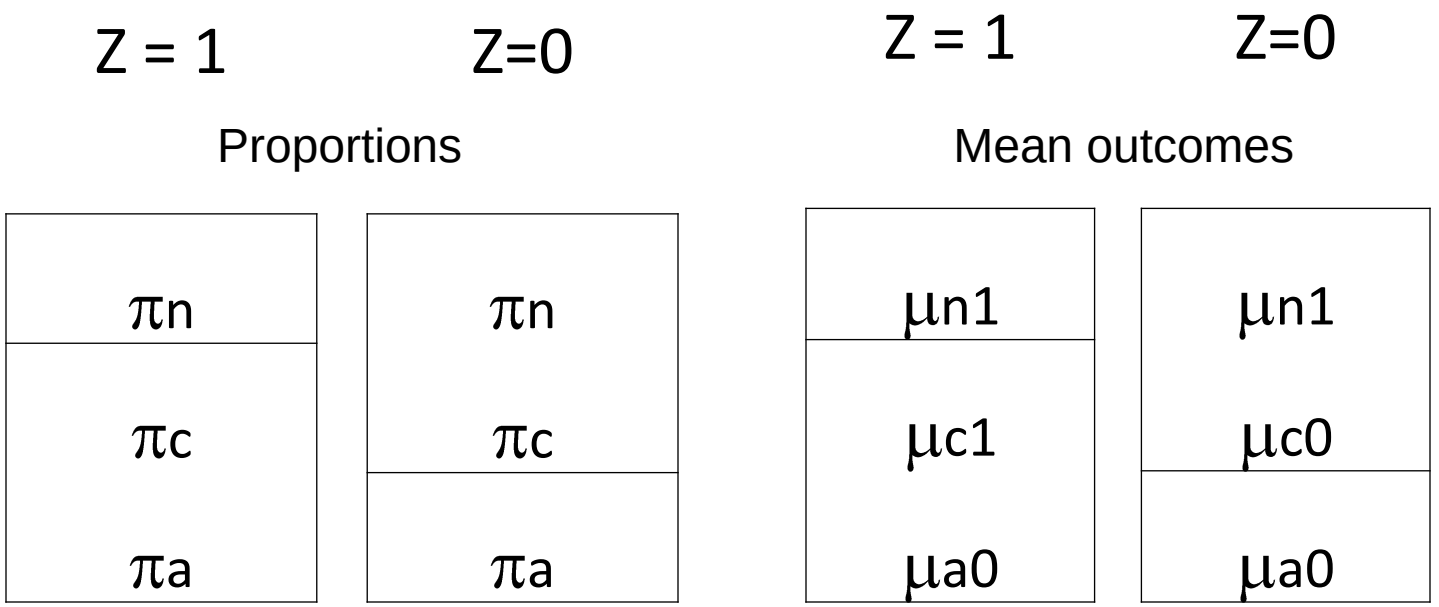
- No effect of treatment assignment for those who would not change their treatment receipt (or mediator) in response to treatment assignment –never-takers and always-takers in the AIR example.
- Treatment assignment itself may have some effect on the outcome (e.g., may have a positive or a negative effect, in particular, on psychological outcomes).
- This effect is less likely in blinded experiments if it is hard for study participants to tell which treatment condition they are in.

Four Compliance Types (AIR 1996)

- Individuals are cross-classified into 4 latent compliance types (*Ci*) based on their *potential* treatment receipt behavior *M* under *both* treatment conditions

$$C_i = \begin{cases} n \text{ (never-taker)} & \text{if } M_i(1) = 0, \text{ and } M_i(0) = 0, \\ d \text{ (defier)} & \text{if } M_i(1) = 0, \text{ and } M_i(0) = 1, \\ c \text{ (complier)} & \text{if } M_i(1) = 1, \text{ and } M_i(0) = 0, \\ a \text{ (always-taker)} & \text{if } M_i(1) = 1, \text{ and } M_i(0) = 1. \end{cases}$$

Identification of CACE



Identification of CACE (AIR 1996)

- Under monotonicity, $\pi_d = 0$. Then, π_c is derived as
  - $\pi_c = 1 - \pi_n - \pi_a$.
- Under exclusion restriction, $\mu_{n1} = \mu_{n0}$ and $\mu_{a1} = \mu_{a0}$
- Then, CACE is derived as

$$\gamma_c = \frac{\mu_1 - \mu_0}{\pi_c} ,$$

where all involved parameters are directly estimable from the observed data (e.g., $\bar{Y}_1$, $\bar{Y}_0$, $\hat{\pi}$)

Identification Strategies

- Use assumptions that are plausible based on science
- Bounding based on possible ranges of parameter values. Very unreasonable assumptions can be useful (we are very clear about the direction of violation).
- Multiple perspectives on the same unknown parameter (Jo & Vinokur, 2011)
  - Exclusion restriction: the mean potential outcome of never-takers under the control is the same as that under the treatment condition.
  - An alternative: the mean potential outcome of never-takers under the control is the same as that of compliers under the same condition.
- Utilize auxiliary information such as from covariates, multivariate (multiple, longitudinal) outcomes, etc.

The Role of Covariates in CACE estimation

- In already well identified models (non-parametrically), covariates that are good predictors of compliance
  - Increase precision (power) in the estimation of CACE (Jo, 2002)
  - Reduce bias due to the exclusion restriction violation (Jo, 2002) and due to deviation from normality (parametric approaches)
- Alternative identifying assumptions can be formulated based on pretreatment covariates
  - Additivity (no treatment*covariates interaction) assumption can replace other identifying assumptions, e.g., the exclusion restriction (Jo, 2002).
  - Sequential ignorability applied to M as used in Imai et al., 2010; Principal ignorability applied to C as used in Jo & Stuart, 2009

Ignorability Assumptions

5. Sequential Ignorability: treatment receipt (M) is independent of the potential outcomes given the observed pretreatment covariates (Imai et al., 2010).
6. Principal Ignorability: compliance type (C) is independent of the potential outcomes given the observed pretreatment covariates (Jo & Stuart, 2009).

- 5 or 6 may replace 3 & 4 (exclusion, monotonicity)
- The relationship between 5 and 6 needs to be further studied – may help better understand the role of compliance; may serve as sensitivity analysis.

Principal Stratification and Propensity Score Methods

- Propensity score based methods mostly rely on ignorability (sequential ignorability), which makes elegant modeling of complex problems possible.
- One advantage of the principal stratification approach is in its open structure – no one set of fixed assumptions (e.g., exclusion+monotonicity, sequential ignorability).
- The two methods complement each other (e.g., Jo, 2008; Jo & Stuart, 2009; Stuart & Jo, 2012)
  - Understanding a particular problem from different perspectives
  - Sensitivity analysis

Example: JOBS II

- Job Search Intervention Study (Vinokur, Price, & Schul, 1995; Vinokur & Schul, 1997).
- A randomized field trial to improve mental health and to promote high-quality reemployment among unemployed workers.
- 54% compliance (attended 1 or more of 5 seminars).
- It turned out that study participants' expected compliance measured at baseline was a strong predictor of actual compliance
  - "How likely or unlikely is it that you would participate in the one-week job seminar if you were offered the opportunity during the next three weeks?"

Estimation of CACE

- When CACE is non-parametrically identified, for example using the exclusion restriction and monotonicity, simpler estimation methods such as the instrumental variable approach (IV) can be used, even with covariates (see Bloom, 1984).
- Maximum-likelihood (ML) using the EM algorithm, e.g., Mplus mixture analysis feature
- Bayesian estimation method (Hirano et al., 2000)
- Propensity score methods (Jo & Stuart, 2009)

Sensitivity Analysis

- The CACE estimates can be sensitive (i.e., biased) to violation of underlying identifying assumptions (monotoncity, exclusion restriction, ignorability, etc).
- Since identification of CACE relies on assumptions that are not verifiable based on observed data, the quality of inference relies on high quality sensitivity analysis.
- This is an essential step in CACE estimation and causal modeling in general.

Plausibility of Identifying Assumptions

- Our sense of relative plausibility can be quite off.
- Seemingly more plausible assumptions may or may not result in less biased estimates (each assumption follows its own bias mechanism).
- Relying on intuitive notions of which assumptions are more or less plausible than others can be dangerous – sensitivity analysis is always important.

Sensitivity Analysis Strategies

- Closely related to identification strategies.
- Utilize alternative assumptions & estimation methods
- Utilize auxiliary information, e.g., proper priors (Hirano et al., 2000) and covariates (Jo, 2002; Jo & Stuart, 2009)
- Construct bounds – utilize science, experts' knowledge
- Easiest ways (better than no sensitivity analysis!)
  - Monitor sensitivity of the CACE estimate to inclusion of covariates in the model
  - Use at least a couple of methods that rely on substantially different sets of identifying assumptions

Handling simultaneous complications

- In practice, several complications co-exist.
- Noncompliance and nonresponse (outcome missingness) - Frangakis & Rubin, 1999; Jo, 2008; Jo et al., 2010; Mealli et al., 2004; Grilli & Mealli, 2008; O'Malley & Normand, 2004; Peng et al., 2004.
- Noncompliance and interference (SUTVA violation) - Frangakis et al., 2002; Jo et al., 2008; VanderWeele, T.J. & Tchetgen Tchetgen, 2011; Sobel/Hudgens & Halloran extensions. This is a topic that needs more investigation.
- Conducting high quality rigorous sensitivity analysis becomes an even more challenging task.

Another challenging topic: multi-strata CACE

- Identification of causal effects becomes very difficult as the number of strata increases (more levels of treatment receipt, multiple comparison groups or treatments)
  - Identification involves more unknown quantities
  - More and stronger assumptions are usually necessary
  - The quality of causal effect estimation is likely to go down
  - Sensitivity analysis becomes very difficult
- Some possibilities
  - Rely on strong predictors of compliance (covariates)
  - Simplify - reduce the number of strata by coarsening (Jin & Rubin, 2008; Jo, Wang, Ialongo, 2009)
  - This is an area that needs much development.

Multiple strata example: 2 active treatment arms

- 3 choices under each condition – receive A, receive B, or receive neither (O).

$$C_i = \begin{cases} 1 \text{ (stratum AA)} & \text{if } M_i(A) = A, \text{ and } M_i(B) = A \\ 2 \text{ (stratum AB)} & \text{if } M_i(A) = A, \text{ and } M_i(B) = B, \\ 3 \text{ (stratum AO)} & \text{if } M_i(A) = A, \text{ and } M_i(B) = O, \\ 4 \text{ (stratum BA)} & \text{if } M_i(A) = B, \text{ and } M_i(B) = A \\ 5 \text{ (stratum BB)} & \text{if } M_i(A) = B, \text{ and } M_i(B) = B, \\ 6 \text{ (stratum BO)} & \text{if } M_i(A) = B, \text{ and } M_i(B) = O, \\ 7 \text{ (stratum OA)} & \text{if } M_i(A) = O, \text{ and } M_i(B) = A \\ 8 \text{ (stratum OB)} & \text{if } M_i(A) = O, \text{ and } M_i(B) = B, \\ 9 \text{ (stratum OO)} & \text{if } M_i(A) = O, \text{ and } M_i(B) = O, \end{cases}$$

- Even after applying monotonicity and the exclusion restriction, the model is still non-identified with 3 different causal effects to identify (AB, AO, OB).

References

- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. Journal of the American Statistical Association, 91, 444–455.
- Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. Evaluation Review, 8, 225–246.
- Frangakis, C. E. & Rubin, D. B. (1999). Addressing complications of intention-to-treat analysis in the presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika, 86*, 365-379.
- Frangakis, C. E. & Rubin, D. B. (2002). Principal stratification in causal inference. Biometrics, 58, 21–29.
- Frangakis, C. E., Rubin, D. B., & Zhou, X. H. (2002). Clustered encouragement design with individual noncompliance: Bayesian inference and application to advance directive forms. Biostatistics, 3, 147–164.
- Grilli, L., & Mealli, F. (2008). Nonparametric bounds on the causal effect of university studies on job opportunities using principal stratification. *Journal of Educational and Behavioral Statistics, 33*, 111-130.
- Hirano, K., Imbens, G. W., Rubin, D. B., & Zhou, X. H. (2000). Assessing the effect of an influenza vaccine in an encouragement design. Biostatistics, 1, 69–88.
- Holland, P. (1986). Statistics and causal inference. Journal of the American Statistical Association 81 (396), 945–960.

References

- Hong, G. and S. Raudenbush (2006). Evaluating kindergarten retention policy. Journal of the American Statistical Association 101 (475), 901–910.
- Hudgens, M. and M. Halloran (2008). Toward causal inference with interference. Journal of the American Statistical Association 103 (482), 832–842.
- Ialongo, N. S., Werthamer, L., Kellam, S. G., Brown, C. H., Wang, S., & Lin, Y. (1999). Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression and antisocial behavior. American Journal of Community Psychology, 27, 599–642.
- Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference, and sensitivity analysis for causal mediation effects. Statistical Science, 25, 51–71.
- Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. Psychological Methods, 15, 309–334.
- Jin, H., & Rubin, D. B. (2008). Principal stratification for causal inference with extended partial compliance. Journal of the American Statistical Association, 103, 101–111.
- Jo, B. (2002). Estimating intervention effects with noncompliance: Alternative model specifications. Journal of Educational and Behavioral Statistics, 27, 385–420.
- Jo, B. (2002). Model misspecification sensitivity analysis in estimating causal effects of interventions with noncompliance. Statistics in Medicine, 21, 3161–3181.

References

- Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. Psychological Methods, 7, 178–193.
- Jo, B., Asparouhov, T., Muthén, B. O., Ialongo, N. S., & Brown, C. H. (2008). Cluster randomized trials with treatment noncompliance. Psychological Methods, 13, 1-18.
- Jo, B. (2008). Bias Mechanisms in intention-to-treat analysis with data subject to treatment noncompliance and missing outcomes. *Journal of Educational and Behavioral Statistics, 33*, 158-185.
- Jo, B., Ginexi, E. M., & Ialongo, N. S. (2010). Handling missing data in randomized experiments with noncompliance. *Prevention Science*, *11,* 384-396.
- Jo, B., & Vinokur, A. (2011). Sensitivity analysis and bounding of causal effects with alternative identifying assumptions. *Journal of Educational and Behavioral Statistics, 36*, 415-440.
- Mealli, F., Imbens, G. W., Ferro, S., & Biggeri, A.(2004). Analyzing a randomized trial on breast self-examination with noncompliance and missing outcomes. *Biostatistics, 5*, 207-222.
- O'Malley, A. J., & Normand, S. L. T. (2004). Likelihood methods for treatment noncompliance and subsequent nonresponse in randomized trials. *Biometrics, 61*, 325-334.

References

- Peng, Y., Little, R. J., & Raghunathan, T. E. (2004). An extended general location model for causal inferences from data subject to noncompliance and missing values. *Biometrics, 60*, 598-607.
- Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. Journal of the American Statistical Association, 101, 1398–1407.
- Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. Annals of Statistics, 6, 34–58.
- Rubin, D. B. (1980). Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by D. Basu. Journal of the American Statistical Association, 75, 591–593.
- Rubin, D. B. (1990). Comment on "Neyman (1923) and causal inference in experiments and observational studies." Statistical Science, 5, 472–480.
- VanderWeele, T.J. & Tchetgen Tchetgen, E.J. (2011). Bounding the infectiousness effect in vaccine trials. Epidemiology, 22, 686-693.
- Vinokur, A. D., Price, R. H., & Schul, Y. (1995). Impact of the JOBS intervention on unemployed workers varying in risk for depression. American Journal of Community Psychology, 23, 39–74.