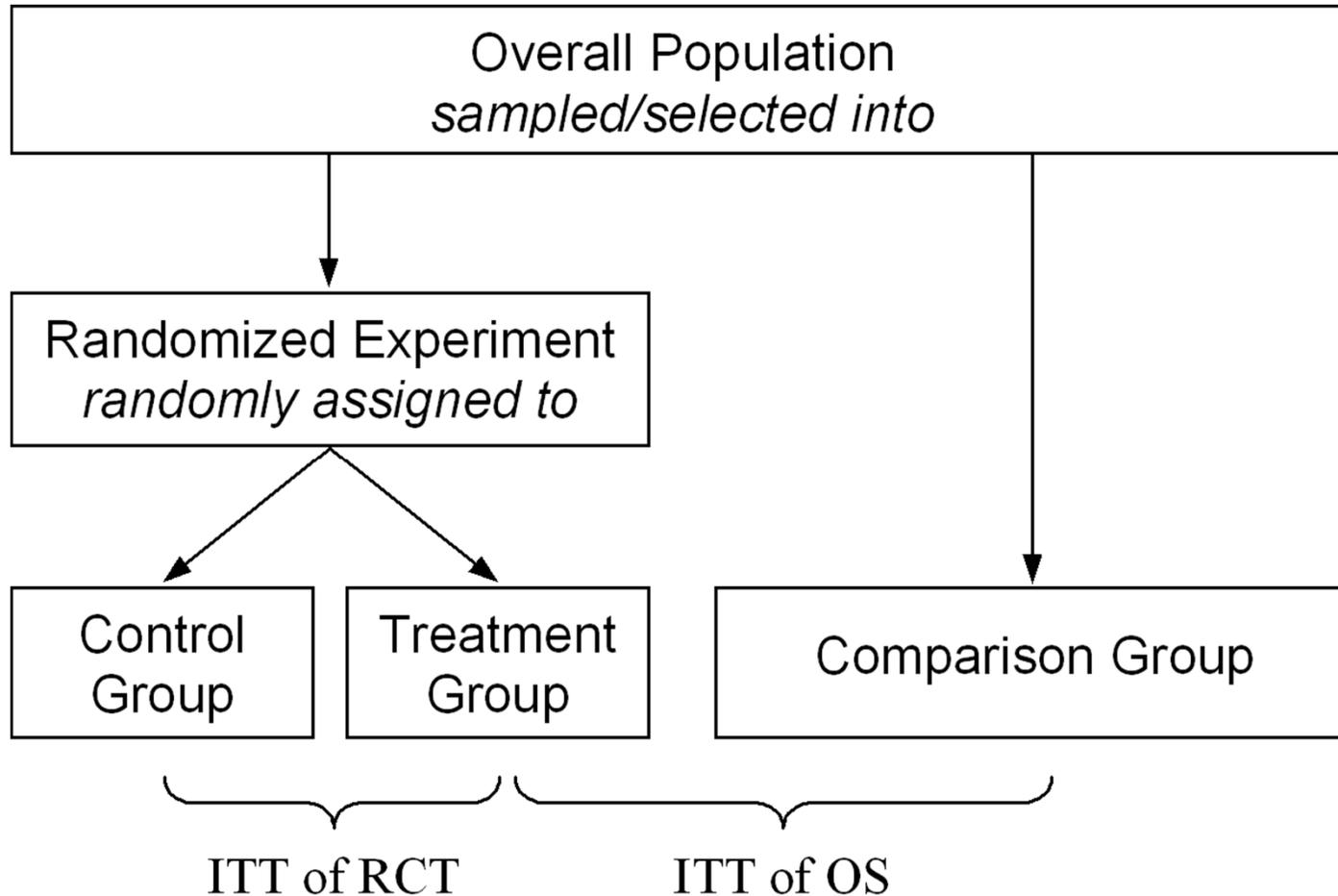


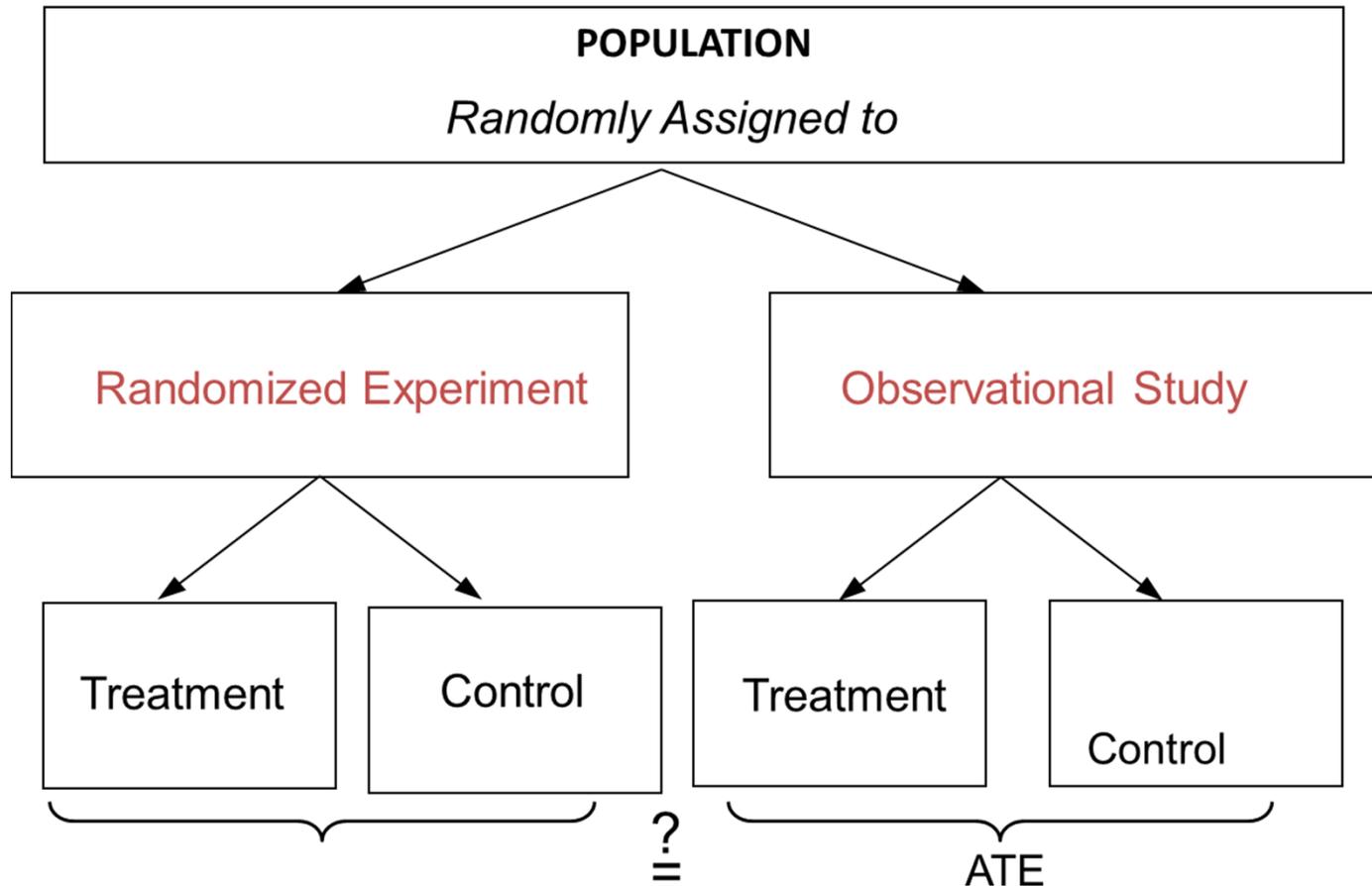
Which Quasi-Experiments Might Provide Presumptively Unbiased Causal Estimates?

Thomas D. Cook
Northwestern University
November 18, 2013

WSC Three-Arm Design



WSC Four-Arm Design



The Issues

- NOT an issue of whether RCT estimates can be recreated. This is theoretically trivial.
- With enough heterogeneous cases it is an issue of inducing how often a given design recreates RCT estimates.
- It is an issue of deducing and testing the conditions under which a given design reproduces better or worse approximations to RCT causal estimates

RCTs and RD

- 7 within-study comparisons in different domains. All report similar causal estimates at the cutoff
- This is theoretically trivial but nice for practice
- One deliberate study of comparison RD functions (CRD) shows the same results as RCT in all the areas away from the cutoff
- What Works Clearinghouse (WWC) standards for RD are now being revised, but the proposed revisions do not deal with the advantages of CRD. This seems odd.

RCT and ITS

- 7 studies comparing RCT and ITS results, all but one CITS
- All show the same results for changes in mean and (where tested) slope
- WWC doesn't include ITS or CITS except in the limited context of single case designs
- This seems short-sighted and should be changed

RCTs and Non-equivocal Control Group Design (NECGD)

- Over 20 studies in many different fields but skewed toward job training and education reform
- Shows close approximation to RCT results with:
 - “Local” comparison group selection;
 - Pretest measures of outcome being used to match groups;
 - Greater independent knowledge and consideration of the likely true selection process;
 - Multi-level matching;
 - Hybrid matching strategy. This combines local matching where it is close with non-local focal matching when local matches do not work

Cautions about RCTs and NECGDs

- Studies show: All of the procedures above have been shown to reduce all of the bias sometimes
- All of the above reduce some of the bias almost always
- None of these can be guaranteed to reduce all bias all of the time, absent complete knowledge of the selection process
- WSC research is moving apace to combine the strategies listed above to improve even more the prediction of when RCT results are stably (but not invariably) replicated

Final Note on RCTs and NECGD

- WWC accepts NECGD only if the treatment and control groups do not vary at pretest on measures of the outcome
- This is naïve since
 - a) it does not deal with time-varying pretest mean differences by group; and
 - b) not all designs have pretest measures but utilize some of the matching techniques above
- More sophisticated consideration needs to be given to WWC standards for accepting NECGDs as adequate

Conclusion

- I agree with most of what is in WWC and generally endorse its standards. However,
 - Greater consideration needs to be given to special status of CRD
 - Standards are needed for accepting CITS designs
- The WWC discussion of NECGDs is out-of-date and needs reconsideration in light of the weight placed on no differences in pretest means between treatment and control group
- Thank you