

RESEARCH METHODS & REPORTING

Using instrumental variables to adjust for treatment contamination in randomised controlled trials

Jeremy B Sussman,^{1,2} Rodney A Hayward^{1,2}

¹Robert Wood Johnson Clinical Scholars Program, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA

²Department of Veterans Affairs, Veterans Affairs Health Services Research and Development Center of Excellence, Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, Michigan, USA

Correspondence to: J B Sussman jeremysu@med.umich.edu

Accepted: 14 November 2009

Cite this as: *BMJ* 2010;340:c2073
doi: 10.1136/bmj.c2073

When patients in a randomised controlled trial do not follow their assigned treatment protocol, the resultant “treatment contamination” can cause misleading findings. **Jeremy B Sussman** and **Rodney A Hayward** describe a simple yet rarely used analytic technique, the “contamination adjusted intention to treat analysis,” which complements the intention to treat approach by producing a better estimate of the benefits and harms of receiving a treatment

Introduction

The recent European Randomized Study of Screening for Prostate Cancer¹ concluded that 1400 patients would need prostate specific antigen screening in order to prevent one death from prostate cancer. This number will be used in meta-analyses, cost effectiveness analyses, and clinical guidelines. But is it accurate? In the study, nearly 20% of people assigned to receive prostate specific antigen screening didn't undergo a single test in 10 years. In a similar American randomised controlled trial,² 40% of participants who weren't supposed to receive prostate specific antigen screening were actually tested. Both these trials were analysed and interpreted as though all participants followed the treatment they were randomised to. If participants had in fact adhered to their assigned intervention as randomised, would the results have been different? Are the results accurate when it comes to advising individual patients whether they should be screened?

Randomised controlled trials are the “gold standard” for examining the efficacy and safety of medical interventions

SUMMARY POINTS

When patients in a clinical trial do not follow the protocol for their assigned treatment, the resultant “treatment contamination” can produce misleading findings

The older methods used to deal with this problem, the “as treated” and “per protocol” analysis techniques, are flawed and inaccurate

Intention to treat analysis estimates the effect of recommending a treatment to study participants, not the effect of receiving the treatment

A technique that we call “contamination adjusted intention to treat analysis” can complement the intention to treat approach by producing a better estimate of the benefits and harms of receiving a treatment

The contamination adjusted intention to treat analysis uses the statistical technique of instrumental variable analysis to address contamination

Glossary

- **As treated:** Method of analysis for randomised controlled trials in which all patients are analysed on the basis of the treatment ultimately received, regardless of the treatment to which they were randomly assigned.
- **Contamination adjusted intention to treat:** Method of analysis for randomised controlled trials in which all patients are analysed as they were randomised and then the result adjusted for treatment contamination by using an instrumental variable.
- **Crossover:** When a study participant receives the intervention for the group to which he or she has not been assigned.
- **Instrumental variable:** An analytical technique, traditionally used in non-randomised research studies, that uses a variable associated with the factor under study but not directly associated with the outcome variable or any potential confounders.
- **Intention to treat:** Method of analysis for randomised controlled trials in which all patients are analysed as they were randomised, regardless of behaviour or treatment received.
- **Loss to follow-up:** When a participant in a study is not involved in the outcome assessment. This issue is not addressed in this paper.
- **Non-adherence:** When a study participant does not receive the assigned therapy, whatever the cause and no matter how legitimate the reason. Non-adherence is often used differently in clinical practice from in experimental research.
- **Per protocol:** Method of analysis for randomised controlled trials in which individuals are included in the analysis only if they followed the assigned protocol and are removed from the analysis entirely if they do not follow protocol.
- **Treatment contamination:** Any time the study participant does not follow the protocol for the assigned treatment. In a study with no treatment contamination, the results of the intention to treat and contamination adjusted intention to treat analyses will be identical.

because they are considered free from bias. However, what was seen in the recent prostate specific antigen trials is common in randomised controlled trials and makes the published data less reliable. When study participants do not receive the treatment to which they were randomised, the flaw called “treatment contamination” is created. Treatment contamination can occur through treatment non-adherence (not receiving the recommended intervention because of treatment intolerance or patient preference) and treatment crossover (receiving the intervention intended for the other group in a trial). Just as non-adherence is common in clinical practice, treatment contamination in randomised controlled trials is

Different methods of analysing a randomised controlled trial

	Explanation	Benefits	Negatives
As treated	Analyses by treatment received	Easy to calculate Tries to address patient oriented question	Results in non-random omission bias
Per protocol	Omits all participants who do not follow protocol	Easy to calculate Tries to address patient oriented question	Results in non-random omission bias
Intention to treat	Analyses by randomisation, ignores whether treatment received	Easy to assess Provides good estimate of the effect of recommending a treatment to study population	Underestimates value of receiving the treatment
Contamination adjusted intention to treat	Analyses by randomisation, adjusts for whether treatment received	Provides good estimate of an individual's risks and benefits of receiving a treatment	Overestimates population level treatment benefits Somewhat more difficult calculation

not a small or infrequent problem—some of our largest trials have contamination of more than 30%.³⁻⁶

In this paper, we describe a method called the “contamination adjusted intention to treat” (CA ITT) analysis that better estimates the benefits of receiving a treatment. CA ITT analysis uses an established statistical technique called instrumental variables (IVs) analysis to adjust for the bias created by contamination. CA ITT analyses could be an excellent complement to traditional analyses but are rarely used in clinical trials, which have traditionally emphasised analytical simplicity.⁷⁻¹⁰ We outline the problem of treatment contamination and how it is currently addressed, describe IVs and CA ITT, how they can be used in clinical trials, and summarise the benefits and limitations of the CA ITT technique.

Background: What do we do now, and what's wrong with it?

How to address treatment contamination in randomised controlled trials has been debated for years. The most commonly used approaches historically are the “as treated” and “per protocol” techniques, which have appropriately fallen out of favour (table). In as treated analyses, participants are analysed entirely on the basis of treatment received (that is, according to their behaviour, not their random assignment). In per protocol analyses, participants who fail to follow the protocol are simply dropped from the study (that is, included or excluded according to their behaviour). Although these approaches seem reasonable, people's behaviours are strongly non-random, with non-adherents generally being less healthy and less health conscious than those who adhere to treatment.¹¹⁻¹² Analysing trial data by behaviour removes the benefit of randomisation, yielding results that are generally biased.¹¹⁻¹³

Guidelines now recommend that randomised controlled trials use intention to treat analysis.¹⁴⁻¹⁵ In intention to treat analysis, sometimes called “analyse as randomised” analysis,¹⁴⁻¹⁵ participants are analysed on the basis of the treatment arm to which they were initially assigned, regardless of their ultimate treatment exposure. Intention to treat analysis avoids the flaws of the as treated and per protocol approaches and should always be the initial analysis of a randomised controlled trial. However, intention to treat analysis ignores treatment contamination altogether. What often goes unappreciated is the fact that intention to treat analysis answers the question: “how much do study participants benefit from being assigned to a treatment group?” This can be impor-

tant information to policy makers and health planners, but patients and clinicians generally want to know the answer to a different question: “what are the risks and benefits of receiving a treatment?”

Using instrumental variables to help us understand results of randomised controlled trials

When physicians were investigating whether maternal smoking leads to poor birth outcomes, they determined that evidence from traditional longitudinal studies was unreliable because smokers and non-smokers are behaviourally different in so many ways aside from smoking status. A solution to this problem was found through tobacco taxes. Throughout the 1970s and 1980s, American states had many and varied increases in tobacco taxes that clearly altered smoking rates across states. Increases in cigarette taxes seem unlikely to affect birth outcomes other than through the effect on smoking. Researchers used a statistical technique known as IV analysis to estimate the impact of smoking on birth outcomes by comparing birth outcomes and changes in smoking rates with state changes in cigarette taxes.¹⁶

IV techniques were designed to learn from natural experiments (that is, changes in people's environments or exposures unrelated to their individual choices or behaviours), in which an unbiased “instrument” (such as change in tobacco taxation level) makes the exposure of interest (such as an individual's smoking habits) more or less likely but has no other effect, either directly or indirectly, on the outcome. IV analysis assesses how the instrument predicts the exposure and the outcome, then uses that information to understand how the exposure predicts the outcome. This type of two stage analysis is used often in the social sciences and is becoming common in observational studies in medicine.¹⁷⁻²⁰ Medical examples of IVs in observational studies include comparing the outcomes among patients whose doctors prefer first generation (conventional) versus second generation (atypical) antipsychotics,¹⁹ analysing the effects of dramatic changes in copayment on medication adherence and subsequent health outcomes among patients with chronic heart failure,²¹ and using distance from a catheterisation lab to approximate the benefit of intensive therapy for heart attacks.¹⁷ Even the random assortment of genetic information has been used as an IV called “Mendelian randomisation”—a genetic variant acts as an instrumental variable to help disentangle the confounded causal relation between phenotype and disease.²²

Despite their theoretical appeal, valid IVs are uncommon. The most common IV in medicine is randomisation within a randomised controlled trial that has treatment contamination. IV analysis can bridge the gap between the more policy focused question posed by intention to treat analyses and the patient focused question of biological efficacy.^{8,23,24}

The concept behind a CA ITT analysis is quite simple. The randomised controlled trial is treated as an IV, with treatment assignment as the “instrument.” The effect of treatment assignment on outcome observed (intention to treat analysis) is adjusted by the percentage of assigned participants who ultimately receive the treatment (contamination adjustment). This way the effect of treatment receipt on the risk of the outcome can be obtained (figure). If a patient stops taking a medication early, the measured

bmj.com archive
Previous Research
Methods and Reporting
articles

- Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses (*BMJ* 2010;340:c117)
- Rethinking pragmatic randomised controlled trials: introducing the “cohort multiple randomised controlled trial” design (*BMJ* 2010;340:c1066)
- CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials (*BMJ* 2010;340:c869)
- CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials (*BMJ* 2010;340:c332)
- Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes (*BMJ* 2010;340:b5087)

non-adherence can be adjusted for with CA ITT and the biases of as treated and per protocol analyses can be avoided. The intention to treat and the CA ITT estimates provide complementary information, and both results are important in their own right.^{7 8 24-28} IV estimators are available in most major statistical packages.

Loss to follow-up, where a patient does not participate in planned follow-up evaluations in a trial or cannot be evaluated at all, is also a major cause of bias in randomised controlled trials, but is a separate and statistically more difficult problem (for a discussion of this topic, see Dunn et al¹⁰).

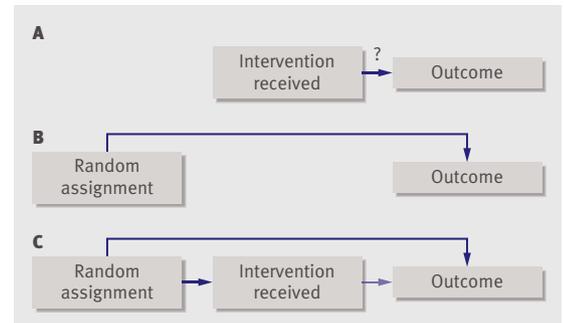
Real world examples

CA ITT analysis has been used for randomised controlled trials in the past, but rarely, and has gone by several different names, including “complier average causal effects,”^{9 23} IVs,^{29 30} efficacy estimator,³¹ and preference based analysis.²⁶ For example, one study found that vitamin A supplementation in malnourished children reduced mortality by 41% when a traditional intention to treat analysis was used.³¹ Once the high mortality rate of the treatment non-adherents was accounted for, however, receiving supplements was found to reduce mortality by two thirds (72%). A recent CA ITT analysis determined that faecal occult blood testing reduces mortality from colorectal cancer by 25% if the patient returns the sample cards, rather than by 19%, the value associated with being randomised to receive the test kit.³²

Adjusting randomised trials for treatment contamination could allow more nuanced understanding of the medical literature. For example, the 1994 Scandinavian Simvastatin Survival Study (4S)³³ found a 20% greater effect of simvastatin on cardiac events than did the 2002 Heart Protection Study (HPS).³⁴ Although other explanations are possible, differences in treatment contamination clearly played a role. The Heart Protection Study, which unlike the Scandinavian Simvastatin Survival Study occurred after statins were in general use, had much higher treatment contamination. This difference alone would account for almost all of the difference between the intention to treat estimates reported in these two trials.

What are the benefits and limitations of CA ITT estimates?

The primary benefit of CA ITT analysis is improved accuracy in estimating the size of treatment benefit for a patient who receives the treatment. Why is this important? The exact effect size of a treatment is relevant whenever you weigh a treatment against negative consequences like side effects, against treatment costs in cost effectiveness analyses, or against another treatment. For example, if many patients stop taking a weight loss drug because of side effects, an intention to treat analysis would measure early treatment side effects appropriately but assess long term weight loss and side effects quite poorly. Patients and clinicians would be better informed by having information on both estimates: the proportion of people who stop the medicine because of side effects; and the degree of weight loss and side effects in patients who continue taking the medication. The comparative effectiveness of two chemotherapeutic regimens with different contamination rates could similarly be misleading to patients and clinicians.^{8 32}



Analysis of randomised controlled trials. In per protocol and as treated analyses (A), random assignment is ignored, creating less reliable results. In intention to treat analyses (B), only the effect of randomisation is assessed, not the effect of receiving the intervention. The two stage contamination adjusted intention to treat approach (C) uses assignment and intervention received to calculate the effect of receiving the treatment

In addition, the estimate of benefit in many cost effectiveness analyses is based on the outcome of an intention to treat analysis, but the estimated treatment cost assumes 100% adherence. This disparity will result in underestimation of the treatment’s cost effectiveness and will particularly disadvantage effective but costly treatments.

CA ITT analysis of randomised controlled trials has some limitations. Treatment contamination is not always assessed in clinical trials and can be quite difficult to measure, especially in trials of medications for chronic conditions. Partial contamination and adherence can complicate analysis substantially. Most importantly, the CA ITT technique assumes that if non-adherents had received the treatment, the treatment would have had the same medical effect as it did in adherents. If non-adherent or crossover patients exhibit a different treatment benefit from adherents, a CA ITT analysis will be biased. This issue increases the need to accurately assess heterogeneity of treatment effect among adherents.³⁵ However, although the assumptions of CA ITT can potentially be violated, the assumptions are more reasonable than those of standard intention to treat analysis—that patients who cross over to active treatment get no benefit and those who fail to receive an assigned intervention benefit as though they did receive it.

Advanced IV analysis

There are more detailed IV adjustment techniques than the CA ITT technique described here that can account for complicating factors such as partial contamination (for example, partial medication adherence or early dropout),^{29 36} time dependent contamination (such as in surgery trials where control group participants have surgery late in the study),²⁹ participants lost to follow-up,³⁷ and non-adherence in randomised equivalency studies (that is, comparing two active treatment arms).²⁶ These techniques, although sometimes significantly more complicated than CA ITT, have a strong theoretical basis and can substantially improve the reliability of results.

Conclusion

Randomised controlled trials in which no control participants cross over to active treatment and all those randomised to active treatment tolerate and adhere to their

assigned treatment do not need to be analysed using CA ITT. However, trials do not often proceed like this and are unlikely to do so more often in the future, so an intention to treat analysis will rarely provide an accurate estimate of treatment benefits for those receiving a treatment.

In major trials, efforts should be taken to limit and quantify treatment contamination with close follow-up of study participants, surveys to assess and optimise adherence, and use of intermediate measures of adherence and effectiveness (such as following change in low density lipoprotein in cholesterol trials that are powered for survival). However, contamination is not always just a nuisance factor—it can demonstrate important factors such as how well patients tolerate treatment side effects.

Once deviations from random assignments occur, regardless of the reason, scientists running trials should examine how treatment contamination has affected their results. Although traditional intention to treat analysis (that is, the effect of recommending the treatment to study participants) should still be used for the primary analysis of a randomised controlled trial, the CA ITT approach is also important because it better estimates the efficacy of the treatment in patients who actually receive it.

The authors thank Michelle Heisler, David Kent, and Joshua Angrist for comments on earlier versions of this work.

Contributors: RAH conceived of the idea for this article and was involved in every stage of writing and editing. JBS was involved in all stages of research review, writing, and editing. Both authors act as guarantors.

Funding: This work was supported in part by the Robert Wood Johnson Clinical Scholars Program, the Department of Veteran Affairs Cooperative Studies Program (CSP #465 FS), and the Measurement Core of the Michigan Diabetes Research & Training Center (NIDDK of The National Institutes of Health [P60 DK-20572]). None of the funders had any role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript.

Competing interests: Both authors have completed the Unified Competing Interest form at www.icmje.org/doi_disclosure.pdf (available on request from the corresponding author) and declare: (1) No financial support for the submitted work from anyone other than their employer; (2) No financial relationships with commercial entities that might have an interest in the submitted work; (3) No spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; (4) No non-financial interests that may be relevant to the submitted work.

Provenance and peer review: Not commissioned; externally peer reviewed.

- Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009;360:1320-8.
- Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009;360:1310-9.
- Calverley PM, Anderson JA, Celli B, Ferguson GT, Jenkins C, Jones PW, et al. Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med* 2007;356:775-89.
- Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *N Engl J Med* 2007;356:2257-70.
- Packer M, Fowler MB, Roecker EB, Coats AJ, Katus HA, Krum H, et al. Effect of carvedilol on the morbidity of patients with severe chronic heart failure: results of the carvedilol prospective randomized cumulative survival (COPERNICUS) study. *Circulation* 2002;106:2194-9.
- Jackson RD, LaCroix AZ, Gass M, Wallace RB, Robbins J, Lewis CE, et al. Calcium plus vitamin D supplementation and the risk of fractures. *N Engl J Med* 2006;354:669-83.
- Rubin DR. Inference and missing data. *Biometrika* 1976;63:581-92.
- Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med* 1997;16:1017-29.
- Dunn G. Estimating the causal effects of treatment. *Epidemiol Psychiatr Soc* 2002;11:206-15.
- Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods. *Stat Methods Med Res* 2005;14:369-95.
- Simpson SH, Eurich DT, Majumdar SR, Padwal RS, Tsuyuki RT, Varney J, et al. A meta-analysis of the association between adherence to drug therapy and mortality. *BMJ* 2006;333:15.
- Granger BB, Swedberg K, Ekman I, Granger CB, Olofsson B, McMurray JJ, et al. Adherence to candesartan and placebo and outcomes in chronic heart failure in the CHARM programme: double-blind, randomised, controlled clinical trial. *Lancet* 2005;366:2005-11.
- Schiffner R, Schiffner-Rohe J, Gerstenhauer M, Hofstadter F, Landthaler M, Stolz W. Differences in efficacy between intention-to-treat and per-protocol analyses for patients with psoriasis vulgaris and atopic dermatitis: clinical and pharmacoeconomic implications. *Br J Dermatol* 2001;144:1154-60.
- Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996;313:570-1.
- Straus SE. *Evidence-based medicine: how to practise and teach EBM*. 3rd ed. Elsevier/Churchill Livingstone, 2005.
- Evans WN, Ringel JS. Can higher cigarette taxes improve birth outcomes? *J Public Econ* 1999;72:135-54.
- McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA* 1994;272:859-66.
- Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. *N Engl J Med* 2008;358:771-83.
- Schneeweiss S, Setoguchi S, Brookhart A, Dormuth C, Wang PS. Risk of death associated with the use of conventional versus atypical antipsychotic drugs among elderly patients. *CMAJ* 2007;176:627-32.
- Hearst N, Newman TB, Hulley SB. Delayed effects of the military draft on mortality. A randomized natural experiment. *N Engl J Med* 1986;314:620-4.
- Cole JA, Norman H, Weatherly LB, Walker AM. Drug copayment and adherence in chronic heart failure: effect on cost and outcomes. *Pharmacotherapy* 2006;26:1157-64.
- Sheehan NA, Didelez V, Burton PR, Tobin MD. Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med* 2008;5:e177.
- Angrist JD, Imbens GW, Rubin DR. Identification of causal effects using instrumental variables. *J Am Stat Assoc* 1996;91:444-55.
- Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;29:722-9.
- Glaziosi PP. Meta-analysis adjusting for compliance: the example of screening for breast cancer. *J Clin Epidemiol* 1992;45:1251-6.
- Walter SD, Guyatt G, Montori VM, Cook R, Prasad K. A new preference-based analysis for randomized trials can estimate treatment acceptability and effect in compliant patients. *J Clin Epidemiol* 2006;59:685-96.
- Bang H, Davis CE. On estimating treatment effects under non-compliance in randomized clinical trials: are intent-to-treat or instrumental variables analyses perfect solutions? *Stat Med* 2007;26:954-64.
- Mealli F, Imbens GW, Ferro S, Biggeri A. Analyzing a randomized trial on breast self-examination with noncompliance and missing outcomes. *Biostatistics* 2004;5:207-22.
- Bond SJ, White IR, Sarah Walker A. Instrumental variables and interactions in the causal analysis of a complex clinical trial. *Stat Med* 2007;26:1473-96.
- Eisenberg D, Quinn BC. Estimating the effect of smoking cessation on weight gain: an instrumental variable approach. *Health Serv Res* 2006;41:2255-66.
- Sommer A, Zeger SL. On estimating efficacy from clinical trials. *Stat Med* 1991;10:45-52.
- Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. *Am J Gastroenterol* 2008;103:1541-9.
- Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994;344:1383-9.
- Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002;360:7-22.
- Basu A, Heckman JJ, Navarro-Lozano S, Urzua S. Use of instrumental variables in the presence of heterogeneity and self-selection: an application to treatments of breast cancer patients. *Health Econ* 2007;16:1133-57.
- Sato T. A method for the analysis of repeated binary outcomes in randomized clinical trials with non-compliance. *Stat Med* 2001;20:2761-74.
- Greenland S, Lanes S, Jara M. Estimating effects from randomized trials with discontinuations: the need for intent-to-treat design and G-estimation. *Clin Trials* 2008;5:5-13.