An IV for the RCT: using instrumental variables to adjust for treatment contamination in randomised controlled trials

Jeremy B Sussman Robert Wood Johnson Foundation clinical scholar, Rodney A Hayward professor of medicine and public health

Abstract

Although the randomised controlled trial is the “gold standard” for studying the efficacy and safety of medical treatments, it is not necessarily free from bias. When patients do not follow the protocol for their assigned treatment, the resultant “treatment contamination” can produce misleading findings. The methods used historically to deal with this problem, the “as treated” and “per protocol” analysis techniques, are flawed and inaccurate. Intention to treat analysis is the solution most often used to analyse randomised controlled trials, but this approach ignores this issue of treatment contamination. Intention to treat analysis estimates the effect of recommending a treatment to study participants, not the effect of the treatment on those study participants who actually received it. In this article, we describe a simple yet rarely used analytical 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. This method uses the statistical technique of instrumental variable analysis to address contamination. We discuss the strengths and limitations of the current methods of addressing treatment contamination and the contamination adjusted intention to treat technique, provide examples of effective uses, and discuss how using estimates generated by contamination adjusted intention to treat analysis can improve clinical decision making and patient care.

Introduction

The recent European Randomized Study of Screening for Prostate Cancer 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, 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 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 not a small or infrequent problem—some of our largest trials have contamination of more than 30%.

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. 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 13 Analyzing trial data by behaviour removes the benefit of randomisation, yielding results that are generally biased. 11 12

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, 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 important 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, 21 analysing the effects of dramatic changes in copayment on medication adherence and subsequent health outcomes among patients with chronic heart failure, 22 and using distance from a catheterisation lab to approximate the benefit of intensive therapy for heart attacks. 23 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. 24

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. 1 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 (fig.). If a patient stops taking a medication early, the measured 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. 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 al).

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,” IVs, efficacy estimator, and preference based analysis. 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. 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.

Adjusting randomised trials for treatment contamination could allow more nuanced understanding of the medical literature. For example, the 1994 Scandinavian Simvastatin Survival Study (4S) found a 20% greater effect of simvastatin on cardiac events than did the 2002 Heart Protection Study (HPS). 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 such as 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 short 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.

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. 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), time dependent contamination (such as in surgery trials where control group participants have surgery late in the study), participants lost to follow-up, and non-adherence in randomised equivalence studies (that is, comparing two active treatment arms). 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 changes 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.
### 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.

---

**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/coi_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.

---

1. 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.
2. Andriole GL, Crawford ED, Grubbs RL, 3rd, Buyys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med 2009;360:1310-9.
3. 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.
4. Weinman KJ, Auricchio 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.
5. Packe F, Bower MB, Roeker EB, Coats AJ, Katsis H, Krum H, et al. Effect of candesartan on the morbidity of patients with severe chronic heart failure: results of the candesartan prospective randomized cumulative survival (COPERNCUS) study. Circulation 2002;106:2194-9.
6. 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.
7. Rubin DR. Inference and missing data. Biometrika 1976;63:581-92.
8. Cuziz J, Edwards R, Segnor N. Adjusting for non-compliance and contamination in randomized clinical trials. Stat Med 1997;16:1017-29.
9. Dunn G. Estimating the causal effects of treatment. Epidemiol Psychiatr Soc 2002;11:206-15.
10. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with non-compliance and loss to follow-up: the role of instrumental variables. Stat Methods Med Res 2005;14:369-95.
11. Simpson SH, Ebrahim S, Padwal RS, Tuohy RT, Varney J, et al. A meta-analysis of the association between adherence to drug therapy and mortality. BMJ 2006;333:15.
12. Granger BB, Swedberg K, Eriksson 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;365:2005-11.
13. Schiffrin R, Schiffrin-Rejo J, Gerstenhauer M, Hotzstuder F, Landthaler M, Stolz W. Differences in efficacy between intention-to-treat and per-protocol analysis for patients with psoriasis vulgaris and atopic dermatitis: clinical and pharmacoeconomic implications. Br J Dermatol 2003;148:1154-60.
| Method                        | Explanation                                                                 | Benefits                                                                 | Negatives                                                                 |
|-------------------------------|-----------------------------------------------------------------------------|--------------------------------------------------------------------------|---------------------------------------------------------------------------|
| As treated                    | Analyses by treatment received                                              | Easy to calculate                                                        | Results in non-random omission bias                                        |
|                               |                                                                             | Tries to address patient oriented question                               |                                                                           |
| Per protocol                  | Omits all participants who do not follow protocol                           | Easy to calculate                                                        | Results in non-random omission bias                                        |
|                               |                                                                             | Tries to address patient oriented question                               |                                                                           |
| Intention to treat            | Analyses by randomisation, ignores whether treatment received               | Easy to assess                                                           | Underestimates value of receiving the treatment                            |
|                               |                                                                             | Provides good estimate of the effect of recommending a treatment to study population |                                                                           |
| Contamination adjusted       | 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                           |
| intention to treat            |                                                                             |                                                                          | Somewhat more difficult calculation                                        |
Fig 1 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.