In randomized controlled trials, the intention-to-treat estimator provides an unbiased estimate of the causal effect of treatment assignment on the outcome. However, patients often want to know what the effect would be if they were to take the treatment as prescribed (the patient-oriented effect), and several researchers have suggested that the more relevant causal effect for this question is the complier average causal effect (CACE), also referred to as the local average treatment effect. Sophisticated approaches to estimating the CACE include Bayesian and frequentist methods for principal stratification, inverse-probability-of-treatment–weighted estimators, and instrumental-variable (IV) analysis. All of these approaches exploit information about adherence to assigned treatment to improve upon the intention-to-treat estimator, but they are rarely used in practice, probably because of their complexity. The IV principal stratification estimator is simple to implement but has had limited use in practice, possibly due to lack of familiarity. Here, we show that the IV principal stratification estimator is a modified per-protocol estimator that should be obtainable from any randomized controlled trial, and we provide a closed form for its robust variance (and its uncertainty). Finally, we illustrate sensitivity analyses we conducted to assess inference in light of potential violations of the exclusion restriction assumption.

as-treated analysis; complier average causal effect; instrumental variables; per-protocol analysis; randomized trials

Abbreviations: AT, always takers; BC, baseline compliers; CACE, complier average causal effect; ITT, intention to treat; IV, instrumental variable; NT, never takers; PS, principal stratification; RCT, randomized controlled trial.

The randomized controlled trial (RCT) remains the “gold standard” study design for estimating the causal effect of assigning individuals to active treatment relative to a reference treatment (e.g., standard care). However, when some subjects do not adhere to their assigned treatment, the intention-to-treat (ITT) estimator typically yields an attenuated estimate of the effect of receiving treatment (1, 2). The individual patient who wishes to decide whether to take the treatment or not may want to know the causal effect of taking treatment, not an average causal effect across persons who do and do not take treatment (3); this difference can be viewed as a patient-oriented effect (4).

Researchers are generally aware of the limitations of the ITT estimator for evaluating the causal effect of receiving treatment (5, 6), and some have restricted their analysis to only those persons who complied with the treatment assigned (per-protocol analysis) or have analyzed subjects based on the treatment received (as-treated analysis). These 2 alternatives essentially convert an RCT into an observational study, and because of resulting confounding they do not necessarily capture the relevant causal effect (7, 8).

Balke and Pearl (9, 10) and Angrist et al. (11) independently defined a causal effect in a subpopulation of adherent subjects. Originally called the local average treatment effect (11), it is now more commonly called the complier average causal effect (CACE). The CACE is defined as the average difference in potential outcomes for the response in a subpopulation of subjects who comply with their assigned treatment. These authors proposed the use of a simple but consistent instrumental-variable (IV) ratio estimator (commonly used to analyze observational studies) for estimating the CACE, under some basic assumptions. Other authors have proposed
more sophisticated estimators, which incorporate more baseline covariate information and other assumptions about the distribution of population quantities (12–14).

The purpose of this paper is to familiarize epidemiologists with different perspectives on the CACE and with some simple methods for calculating it and its variance. We present 2 interpretations for the CACE based on formulae that have appeared in various forms in the statistical and epidemiology literature but are not routinely used. We show that the IV ratio estimator for the CACE is better conceptualized as an adjustment to the per-protocol estimator and can be computed with a hand calculator. Its variance depends only on basic summary statistics that should be reported for any RCT with nonadherence. Finally, we provide a simple sensitivity analysis for violations of the exclusion restriction condition.

BACKGROUND

We make the same basic assumptions as those previously described for IV analyses (15). We assume there are 2 levels of treatment, a reference level (control status) and an experimental level (active treatment). If subject $i$ is randomly drawn from the population, we express the causal effect of treatment assignment as $E[Y^{(1)}_i - Y^{(0)}_i]$. Let $Y^{(0)}_i$ and $Y^{(1)}_i$ represent the potential outcomes of the response variable for subject $i$, where $z$ is defined as the level of treatment assignment for subject $i$; $z = 0$ if subject $i$ is assigned to the control status and $z = 1$ if assigned to active treatment. Via the consistency assumption, $Y^{(0)}_i$ is the observed value of the response variable if subject $i$ is assigned to control status and $Y^{(1)}_i$ is the observed value if subject $i$ is assigned to active treatment. We also invoke the common stable unit treatment value assumption that the potential outcomes for subject $i$ are neither influenced by nor influence the potential outcomes for any other subjects and that multiple versions of treatment do not exist. Further assumptions (monotonicity, the exclusion restriction) will be explained as the underlying concepts become relevant.

Franakis and Rubin (12) proposed principal stratification (PS), an approach that characterizes population causal effects conditional on strata defined by baseline characteristics, which are partially identified using posttreatment variables. A principal stratification can be defined with respect to the set of potential outcomes for any posttreatment variable. We define the principal strata with respect to potential outcomes for treatment received ($A$) as $A^{(z)}_i$, where a value of 1 denotes receiving active treatment and a value of 0 denotes being assigned to the control group. As above, $z = 0$ corresponds to control assignment and $z = 1$ corresponds to active treatment assignment. In a participant assigned to active treatment ($z = 1$) and receiving active treatment, $A^{(1)}_i = 1$; in a participant assigned to active treatment and receiving control status, $A^{(1)}_i = 0$. Our results require that the stable unit treatment value assumption must hold for the potential outcomes for both treatment received $A_i$ and outcome $Y_i$.

Within this framework, we obtain 4 possible principal strata (baseline compliers (BC), always takers (AT), never takers (NT), and defiers) for each subject, listed in Table 1 (11). We define 3 population-based parameters for each stratum $s$: the mean values of the potential outcomes for response under active treatment assignment ($\mu^{(1)}_B$) and control assignment ($\mu^{(0)}_B$) and the proportion of the population within the stratum. For reasons of identifiability, we make the common monotonicity assumption, which assumes that no one will take precisely the opposite of his or her treatment assignment; that is, we only allow $s$ to equal BC, AT, or NT (11). Finally, a subject’s PS membership is a baseline characteristic that is not affected by treatment assignment or treatment received (4, 12). Note also that the potential outcomes both for the response and for receiving treatment are defined in terms of treatment assignment. This approach differs from that of Angrist et al., who characterized potential outcomes for response in terms of treatment received (11).

Under the specified PS and the assumptions listed above, the principal effect represents the difference between potential outcomes for participants assigned to treatment in 1 principal stratum (12). We are interested in the average of potential outcomes for the response within the stratum ($S_i$) of baseline compliers, which can be expressed as $E[Y^{(1)}_i - Y^{(0)}_i | S_i = BC] = \mu^{(1)}_B - \mu^{(0)}_B$, where $\mu^{(z)}_B$ is the mean value of the potential outcomes for treatment assignment $z$ in the BC stratum. In other words, the CACE is the causal effect of treatment assignment among the subpopulation of subjects for whom the potential outcomes for level of treatment assignment coincide with the potential outcomes for level of treatment received.

THE IV-PS ESTIMATOR

Under randomized treatment assignment and the stable unit treatment value assumption, the expectation of the ITT estimator can be written as a weighted average of the stratum-specific mean differences (13, 16, 17),

$$E[\bar{Y}_{Z=1} - \bar{Y}_{Z=0}] = \mu^{(1)} - \mu^{(0)} = p_{BC}(\mu^{(1)}_B - \mu^{(0)}_B) + p_{AT}(\mu^{(1)}_A - \mu^{(0)}_A) + p_{NT}(\mu^{(1)}_N - \mu^{(0)}_N), \tag{1}$$

where $\bar{Y}_{Z=1}$ and $\bar{Y}_{Z=0}$ are the sample means among persons assigned to active treatment and control status, respectively, $\mu^{(1)}$ and $\mu^{(0)}$ are the population mean potential outcomes for response among those assigned to active treatment and control status, respectively, and $p_{BC}$, $p_{AT}$, and $p_{NT}$ represent the prevalence of each stratum. Because CACE is equal to $(\mu^{(1)} - \mu^{(0)})$, the ITT estimator will only be unbiased for the CACE when the proportion of baseline compliers in the population is equal to 1 ($p_{BC} = 1$) or the effect is identical in all 3 strata.
Throughout most of this paper, we also assume the exclusion restriction—that is, that there is no direct causal effect of treatment assignment on the response. We can express the exclusion restriction in terms of the mean values of the potential outcomes in the AT and NT strata. If treatment assignment only affects the outcome through treatment received, then both \( \mu_{AT}^{(1)} - \mu_{AT}^{(0)} \) and \( \mu_{NT}^{(1)} - \mu_{NT}^{(0)} \) must equal 0, because the treatment received by always takers and never takers is the same regardless of treatment assignment. If the exclusion restriction holds, then one can identify all relevant population quantities from the observed data. If the exclusion restriction is violated, without additional assumptions about the range of the response variable \((10, 18)\), all that can be done is to assess the sensitivity of the conclusions to different values of these differences. The exclusion restriction may or may not be plausible, depending on the type of treatment assigned and the mechanism of treatment assignment. An example of a potential violation of the exclusion restriction assumption is an RCT on the effects of stretching before exercise to prevent injury \((19)\). Some participants assigned to the stretching program may have adopted additional behaviors, such as extended warm-up, that were not considered part of the intervention. Because warming up prior to exercise is believed to prevent injury, assignment to treatment would have affected the outcome independently of the stretching treatment taken.

Under the exclusion restriction, equation 1 simplifies to \( E[\hat{Y}_{Z=1} - \hat{Y}_{Z=0}] = p_{BC}(\mu_{BC}^{(1)} - \mu_{BC}^{(0)}) \), because the causal effects of treatment assignment in the other 2 strata become zero. Dividing both sides by \( p_{AT} + p_{NT} \) and replacing the population quantities with sample quantities yields the basic IV estimator of Angrist et al. \((11)\),

\[
\hat{Q}_{IV} = \frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{1 - (p_{AT} + p_{NT})}.
\]

The IV-PS estimator, \( \hat{Q}_{IV} \), is a consistent estimator of the CACE under the stated assumptions and is quite simple to calculate, which probably explains its popularity in the statistical and econometric literatures. The variance of \( \hat{Q}_{IV} \) has typically been estimated using the delta method \((20)\) or via the bootstrap \((21)\).

Nonetheless, despite its simplicity and general acceptance in the statistical literature, \( \hat{Q}_{IV} \) still does not routinely appear in analyses of RCTs with nonadherence. There are several possible reasons for this. First, a researcher evaluating policies may only be interested in the ITT causal effect of offering treatment \((4)\). The fact that 49 of 98 trials published in 2008 in \textit{BMJ}, the \textit{New England Journal of Medicine}, \textit{JAMA}, and the \textit{Lancet} included per-protocol or as-treated analyses designed specifically to address nonadherence issues \((22)\) strongly suggests this is not the case. Second, there could be widespread ignorance in the RCT community about the estimator, although this is unlikely given the interest in IV methods in other areas. Third, researchers may have concerns because it seems that important information available in the data is being ignored when computing \( \hat{Q}_{IV} \). Finally, investigators may not want to rely on the exclusion restriction or monotonicity assumptions. In the next 2 sections, we present a different formulation of the IV-PS estimator and a new sensitivity analysis, which address all but the first concern.

**RELATIONSHIP OF THE IV-PS ESTIMATOR TO THE PER-PROTOCOL ESTIMATOR**

The standard per-protocol estimator is the difference between the sample mean for those subjects who were assigned and received active treatment \((\bar{Y}_{Z=1, A=1})\) and the sample mean for those subjects who were assigned and received control status \((\bar{Y}_{Z=0, A=0})\). Using methods similar to the ITT case and assuming that the exclusion restriction holds, one can express the expectation of the per-protocol estimator \((13, 16)\) as

\[
E[\bar{Y}_{Z=1, A=1} - \bar{Y}_{Z=0, A=0}] = p_{BC,A=1}\mu_{BC}^{(1)} - p_{BC,A=0}\mu_{BC}^{(0)} + (1 - p_{BC,A=1})\mu_{AT}^{(0)} + (1 - p_{BC,A=0})\mu_{NT}^{(1)},
\]

where \( p_{BC,A=1} = p_{BC}/(p_{BC} + p_{AT}) \) and \( p_{BC,A=0} = p_{BC}/(p_{BC} + p_{NT}) \).

All quantities other than \( \mu_{BC}^{(1)} \) and \( \mu_{BC}^{(0)} \) in equation 2 can be estimated unbiasedly from the observed data. Under the randomization assumption, we can unbiasedly estimate \( \mu_{AT}^{(0)} \) with \( \bar{Y}_{A=1, Z=0} \), the sample mean of subjects assigned to control status but receiving active treatment. These subjects are \textit{known} to be in the AT stratum, because under monotonicity, members of the AT stratum are the only subjects who will receive active treatment when assigned to control status. Similarily, we can estimate \( \mu_{NT}^{(1)} \) unbiasedly with \( \bar{Y}_{A=0, Z=1} \), the sample mean of those subjects assigned to active treatment but receiving the control designation. Again solving for \( \mu_{BC}^{(1)} - \mu_{BC}^{(0)} \) and replacing population means with sample estimates, we obtain a different formulation for the IV-PS estimator:

\[
\hat{Q}_{IV} = \left( \frac{\bar{Y}_{Z=1, A=1}}{\hat{p}_{BC,A=1}} - \frac{1 - \hat{p}_{BC,A=1}}{\hat{p}_{BC,A=1}} \bar{Y}_{Z=0, A=1} \right) - \left( \frac{\bar{Y}_{Z=0, A=0}}{\hat{p}_{BC,A=0}} - \frac{1 - \hat{p}_{BC,A=0}}{\hat{p}_{BC,A=0}} \bar{Y}_{Z=1, A=0} \right),
\]

where \( \hat{p}_{BC,A=1} = \hat{p}_{BC}/(\hat{p}_{BC} + \hat{p}_{AT}) \) and \( \hat{p}_{BC,A=0} = \hat{p}_{BC}/(\hat{p}_{BC} + \hat{p}_{NT}) \). This is equivalent to the estimator of Little and Yau \((23)\). More importantly, this estimator is exactly equal to that derived in equation 2 \((24)\). Therefore, although the IV-PS estimator \( \hat{Q}_{IV} \) is generally thought of as an adjustment to the ITT estimator, it can equivalently be thought of as an adjustment to the standard per-protocol estimator.

Consider the following example from an RCT in sports medicine \((25)\). Investigators randomized subjects with low back pain to a physiotherapist-supervised exercise regimen.
(the McKenzie method) or standard care, and they recorded the change in pain (via a 10-point numerical rating scale) after 1 week. The control group had no access to active treatment \((p_{AT} = 0)\), and some subjects in the active treatment group did not complete the first week of sessions. Table 2 shows the sample sizes, sample means and standard deviations, and observed adherence for the 4 cross-classification subgroups, as well as values for the 4 estimators.

Note that the per-protocol estimate \(\bar{Y}_{[Z=1, A=1]} = 3.26\) for the mean value of baseline compliers assigned to treatment \((\mu_{BC}^{(1)})\) is equal to the IV-PS estimate of the same quantity because there are no always takers in the study (i.e., \(p_{BC, A=1} = 1\)). However, the naive per-protocol estimate of \(\mu_{BC}^{(0)}\) \(\bar{Y}_{[Z=0, A=0]} = 3.68\) is different from the IV-PS estimate,

\[
\bar{Y}_{[Z=0, A=0]} - \frac{1 - \hat{p}_{BC, A=0}}{\hat{p}_{BC, A=0}} \bar{Y}_{[Z=1, A=0]} = 3.50,
\]

because of the presence of never takers in the active treatment arm. Therefore, the estimated pain reduction for baseline compliers using the consistent IV-PS estimator \((3.26–3.50 = -0.24)\) is less than the potentially biased per-protocol estimate \((3.26–3.68 = -0.54)\).

**APPROXIMATING THE VARIANCE OF THE IV-PS ESTIMATOR**

We have provided an R statistical package (R Foundation for Statistical Computing, Vienna, Austria) as well as an Excel spreadsheet (Microsoft Corporation, Redmond, Washington) that allows for anyone to enter the sample mean, sample standard deviation, and sample size for each of the 4 cross-classification subgroups and obtain both the IV-PS estimate of the CACE and its robust variance estimate. With these basic summary statistics, the robust variance estimator for \(Q_{IV}\) can be written as

\[
V(\hat{Q}_{IV}) = \left(\frac{1}{\hat{n}_{BC, A=1}}\right)^2 F_{11} - 2 \left(\frac{1}{\hat{n}_{BC, A=0}\hat{n}_{BC, A=1}}\right) F_{12} + \left(\frac{1}{\hat{n}_{BC, A=0}}\right)^2 F_{22}.
\]

Conceptually, \(\hat{n}_{BC, A=0}\) is the estimated number of baseline compliers who receive control status, \(\hat{n}_{BC, A=1}\) is the estimated number of baseline compliers who receive treatment, and \(F_{11}, F_{12}, \text{and } F_{22}\) primarily depend on 2 sources of variability: 1) the observed variance within each of the 4 cross-classification subgroups and 2) the discrepancy between the per-protocol and IV-PS estimators. A full explanation of the formula, the necessary assumptions, and its derivation are provided in the Web Appendix (available at http://aje.oxfordjournals.org/).

Given the relative ease with which these quantities can be computed fromRCT data, we recommend that investigators conducting RCTs with nonadherence should report these basic summary statistics to allow computation of \(\hat{Q}_{IV}\). For example, this would allow investigators interested in the CACE to conduct future meta-analyses from RCTs with nonadherence, at almost no additional cost to researchers or journals. Additionally, this simple approach should serve as a baseline analysis for comparisons with more complex analyses (such as G-estimation or IV analysis with covariates).

**SIMULATION RESULTS FOR THE ROBUST VARIANCE ESTIMATOR**

We use a simulation model based on the proposed common-cause principal stratification model of Shrier et al. (4). For each subject, we generate an independent binary random treatment assignment \(Z\) and a common cause of response and treatment received, \(C\), with a standard normal distribution. Conditional on the common cause, we generate potential outcomes for receiving treatment \((A^{(0)}, A^{(1)})\) and potential outcomes for response \((Y^{(0)}, Y^{(1)})\) under 3 different scenarios (described in the Web Appendix and summarized in Web Table 1).

<table>
<thead>
<tr>
<th>Stratum ((S))</th>
<th>Adherence ((A_{(0)}, A_{(1)}))</th>
<th>Stratum Mean Values</th>
<th>Stratum Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline compliers</td>
<td>(A_{(0)}^{(0)} = 0, A_{(1)}^{(1)} = 1)</td>
<td>(\mu_{BC}^{(0)}, \mu_{BC}^{(1)})</td>
<td>(p_{BC})</td>
</tr>
<tr>
<td>Always takers</td>
<td>(A_{(0)}^{(0)} = 1, A_{(1)}^{(1)} = 1)</td>
<td>(\mu_{AT}^{(0)}, \mu_{AT}^{(1)})</td>
<td>(p_{AT})</td>
</tr>
<tr>
<td>Never takers</td>
<td>(A_{(0)}^{(0)} = 0, A_{(1)}^{(1)} = 0)</td>
<td>(\mu_{NT}^{(0)}, \mu_{NT}^{(1)})</td>
<td>(p_{NT})</td>
</tr>
<tr>
<td>Defiers</td>
<td>(A_{(0)}^{(0)} = 0, A_{(1)}^{(1)} = 0)</td>
<td>Assumed not to exist</td>
<td>(\text{NT}), never takers</td>
</tr>
</tbody>
</table>

Abbreviations: AT, always takers; BC, baseline compliers; NT, never takers.

Table 2. Data Summaries for a Randomized Trial Comparing the McKenzie Method With Standard Treatment for Low Back Paina

<table>
<thead>
<tr>
<th>Treatment Assignment</th>
<th>Mean Estimate, points (SD)</th>
<th>Subjects Not Adhering to Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adherent Subjects</td>
<td>Nonadherent Subjects</td>
</tr>
<tr>
<td>Active treatment (n = 70)</td>
<td>3.26 (2.19)</td>
<td>4.16 (2.57)</td>
</tr>
<tr>
<td>Control status (n = 69)</td>
<td>3.68 (2.48)</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; NA, not applicable; SD, standard deviation.

a Summary statistics obtained from the randomized controlled trial by Machado et al. (25) comparing the McKenzie method with standard treatment. Mean values, SDs, and effect estimates are given in points on a 10-point numerical rating scale. Estimates for the different parameters of interest: intention-to-treat: −0.18 (95% CI: −1.04, 0.68); per-protocol: −0.42 (95% CI: −1.27, 0.42); as-treated: −0.52 (95% CI: −1.38, 0.33); instrumental variable–principal stratification: −0.24 (95% CI: −1.32, 0.84).

Figure 1. Results from the first simulated scenario with strong confounding (a common cause has a strong effect on both adherence and the response). A) Distribution (mean value, 25th and 75th percentiles, and 2.5th and 95th percentiles) of the 5 complier average causal effect (CACE) estimators (true principal strata (PS), instrumental variable–principal stratification (IV-PS), intention-to-treat (ITT), per-protocol (PP), and as-treated (AS)). The solid horizontal line represents the true value for the CACE. B) Contribution of the variance and bias towards the overall mean squared error for each of the 5 estimators. The light gray portion of each bar shows the contribution of the variance, and the dark gray portion shows the contribution of the squared bias. C) Distribution of the robust standard error estimates of Abadie (26). The dashed line represents the observed standard deviation of the IV-PS estimator over 1,000 simulated data sets. D) Robust standard error estimates plotted against the estimated proportion of baseline compliers (ρBC).
estimator still provides better coverage and approximately the same mean squared error in estimation. Second, the IV-PS estimator performs fairly well, even relative to the unobtainable true principal stratum memberships. Finally, the robust variance estimates seem to be well calibrated for the scenarios we have considered here and should prove useful in general practice. Larger sample sizes (simulations not shown) exacerbate the inadequacy of the ITT, per-protocol, and as-treated analyses and show that IV-PS would be even more strongly preferred.

SENSITIVITY ANALYSIS FOR THE EXCLUSION RESTRICTION

Under the exclusion restriction, \( \mu_{AT}^{(1)} = \mu_{AT}^{(0)} \) and \( \mu_{NT}^{(1)} = \mu_{NT}^{(0)} \), because there is no direct causal effect of treatment assignment on the outcome. However, this assumption may be unlikely to hold in some contexts (e.g., unblinded studies, behavioral interventions). Although other sensitivity analyses exist in the literature (27, 28), we present a simple method that requires only a minor modification of the IV-PS estimator and its robust variance.

Recall the example of exercise for back pain (Table 2), where the IV-PS estimate (under the exclusion restriction assumption) was \(-0.24\) (95% confidence interval: \([-1.32, 0.84]\)), the ITT estimate was closer to the null, and the per-protocol and as-treated estimates were further from the null. Subjects assigned to exercise who did not complete the first week of sessions (nonadherers) may have received some benefit, violating the exclusion restriction assumption (\( \mu_{NT}^{(1)} \neq \mu_{NT}^{(0)} \), or \( \mu_{NT}^{(0)} - \mu_{NT}^{(1)} \neq 0 \)). Although we cannot differentiate

---

Figure 2. Results from the second simulated scenario with a weak causal effect of a common cause on the response. A) Distribution (mean value, 25th and 75th percentiles, and 2.5th and 95th percentiles) of the 5 complier average causal effect (CACE) estimators (true principal strata (PS), instrumental variable–principal stratification (IV-PS), intention-to-treat (ITT), per-protocol (PP), and as-treated (AS)). The solid horizontal line represents the true value for the CACE. B) Contribution of the variance and bias towards the overall mean squared error for each of the 5 estimators. The light gray portion of each bar shows the contribution of the variance, and the dark gray portion shows the contribution of the squared bias. C) Distribution of the robust standard error estimates of Abadie (26). The dashed line represents the observed standard deviation of the IV-PS estimator over 1,000 simulated data sets. D) Robust standard error estimates plotted against the estimated proportion of baseline compliers (\( \hat{p}_{BC} \)).
μ\textsubscript{NT} from μ\textsubscript{BC} from the data without the exclusion restriction, we can treat this difference as a sensitivity parameter and modify equation 4 to add values for μ\textsubscript{NT}/C\textsubscript{0} and μ\textsubscript{NT}/C\textsubscript{1}:

\[
\hat{Q}_\text{IV} = \frac{Y_{[Z=1,A=1]} - (1 - \hat{p}_{BC,A=1}) (Y_{[Z=0,A=1]} + \mu_{AT} - \mu_{AT}^{(0)})}{\hat{p}_{BC,A=1}}
\]

\[
\hat{Q}_\text{IV} = \frac{Y_{[Z=0,A=0]} - (1 - \hat{p}_{BC,A=0}) (Y_{[Z=1,A=0]} + \mu_{NT} - \mu_{NT}^{(0)})}{\hat{p}_{BC,A=0}},
\]

(6)

The sensitivity analysis for this trial (Figure 4) suggests that one would need to hypothesize a difference of at least −2 in mean pain scores between never takers assigned to usual care and never takers assigned to exercise before the CACE would be statistically significantly different from zero. This magnitude is highly unlikely given that it is twice the difference among any of the 3 observed cross-classification subgroups. Therefore, we believe the observed data do not justify rejecting the null hypothesis even if the exclusion restriction is violated.
It should be noted, however, that such analyses are often considerably flawed.” Our simulations demonstrate support for past work that the per-protocol estimator is not an acceptable surrogate for the IV-PS estimator. Given the ease of calculating the IV-PS estimator and its confidence interval, there is no valid excuse for not reporting them instead.

More advanced estimators have been constructed for estimating the CACE beyond the basic IV-PS estimator. In particular, there exist consistent, adherence-based analyses using additional covariates via instrumental variables (26, 30), inverse probability weighting (8), and G-estimation of structural nested models (24). Adding covariates may improve the analysis in 2 important ways. First, it may reduce the overall uncertainty in the association of the treatment with the response variable by removing known sources of variability. Second, with additional assumptions, one can reduce the uncertainty surrounding stratum membership. However, these more complicated estimators have seen even less general usage in biomedical studies than the IV-PS estimator, probably because of difficulty in both conceptualization and implementation. Finally, clinicians, policy-makers, and meta-analysts interested in patient-oriented effects would not be able to compute them without access to the original data.

Other authors have also proposed sensitivity analyses for PS estimators. First, Egleston et al. (14) present a more complex sensitivity analysis that allows for the use of baseline covariates and does not require the exclusion restriction to hold. In the case of no baseline covariates, their proposed sensitivity analyses would only utilize sample means from the 2 cross-classification groups observed to adhere to treatment and the proportions of those who did not adhere to treatment assignment. In contrast, our proposed per-protocol sensitivity analysis uses all of the available information (means and proportions of all 4 subclassifications) in order to assess sensitivity to violations of the exclusion restriction. Second, McNamee (31) presents a model-based sensitivity analysis that assumes a bivariate normal distribution for potential outcomes and then assesses the bias of the ITT estimator relative to the as-treated and per-protocol estimators for differing proportions of baseline compliers in each treatment arm among those whose status is unknown (rather than the rate in the population). Our approach is simpler and does not require bivariate normality of the potential outcomes. Finally, Jo and Vinokur (32) establish bounds for the IV-PS estimator with the additional complication of nonresponse—for example, with patients who are missing outcome data because they dropped out after the start of the trial. Our robust variance computation could be used to provide confidence limits for those bounds.

In the future, we will investigate extending this work to the case of a binary response variable. Abadie (26) provides robust variance estimators for the binary response case as well, but the computations are not as straightforward. Schwartz et al. (33) provide a parametric model-based sensitivity analysis for use in the case of binary outcomes but also allow for the exclusion restriction to not hold. It may be useful to extend our results to provide closed-form confidence bounds for their sensitivity analyses.

Although we restrict our discussion to RCTs with imperfect adherence to treatment assignment, the RCT is simply an example of a study in which there is a valid instrument and a

**DISCUSSION**

We have argued that the IV-PS estimator should be reported instead of the per-protocol and as-treated estimators for randomized trials with noncompliance. We have also shown how to construct confidence intervals for the CACE from basic summary statistics and how to adjust the estimator and intervals to assess the sensitivity of conclusions to violations of the exclusion restriction. Our work would be applicable in any observational study where a plausibly valid instrument could be used to obtain an estimate of the CACE and should be used as the baseline when trying to incorporate covariates in more complex analyses.

Although the IV-PS estimator is generally thought of as an adjustment to the ITT estimator, considering it an adjustment to the per-protocol estimator provides 3 advantages. First, it is more apparent that only basic per-protocol information is used to compute the IV-PS estimator. Second, it allows for closed-form computation of the robust variance from basic summary statistics, which should be reported from any RCT with nonadherence. Finally, it allows for straightforward sensitivity analyses for violations of the exclusion restriction.

The Explanation and Elaboration for CONSORT 2010 (29) mentions nonadherence twice. The explanation for item 7a (29, p. 8) reads, “Details should be given of any allowance made for attrition or non-compliance during the study.” Our work indicates exactly which details: In all trials, investigators should report the summary statistics for the 4 cross-classification subgroups. Second, the explanation for item 16 (29, p. 18) reads, “Non-compliance with assigned therapy may mean that the intention-to-treat analysis underestimates the potential benefit of the treatment, and additional analyses, such as a per-protocol analysis, may therefore be considered.

![Figure 4. Results from a sensitivity analysis assessing the influence of violation of the exclusion restriction on the exercise trial of Machado et al. (25), conducted in Sydney, Australia, from September 2005 to June 2008. The solid line shows the instrumental variable–principal stratification (IV-PS) estimate of the complier average causal effect (CACE) as a function of the difference between mean counterfactuals in the never taker (NT) stratum. The 2 dotted lines delineate the 95% confidence interval, computed from the robust variance formula. The point and error bars show the estimate and 95% confidence interval for the IV-PS estimate under the exclusion restriction.](image-url)
second variable that also affects the uptake of treatment given the presence/absence of the valid instrument. The methods are therefore applicable to observational studies, but in such studies one would not normally use the terms “per-protocol,” “as-treated,” “baseline compliers,” “always takers,” and “never takers” as in the current paper.

Finally, we emphasize that the IV-PS estimator does not remove other potential problems associated with all adherence-based analyses. First, assigning a binary value for adherence move other potential problems associated with all adherence-they can be widely used for both past and present studies. That should be reported in all RCTs with nonadherence, and thus researchers. We have also provided software for computing a tor uses the same available information yet provides less biased lemm with the CACE itself, not with the IV-PS estimator.

In conclusion, we have demonstrated that the IV-PS estimator uses the same available information yet provides less biased estimates than the ITT, per-protocol, and as-treated estimators, while also being just as easy to compute. We have shown that the IV-PS estimator can be viewed as an adjustment to the per-protocol estimator, which should make it more attractive to researchers. We have also provided software for computing a robust variance estimator using only basic summary statistics and a sensitivity analysis for violations of the exclusion restriction. These simple estimators can be computed easily from data that should be reported in all RCTs with nonadherence, and thus they can be widely used for both past and present studies.

REFERENCES

5. Giorgi Rossi P. Screening: the information individuals need to support their decision: per protocol analysis is better than intention-to-treat analysis at quantifying potential benefits and harms of screening. BMC Med Ethics. 2014;15:28.

ACKNOWLEDGMENTS

Author affiliations: Department of Mathematics and Statistics, Faculty of Science, McGill University, Montreal, Quebec, Canada (Russell J. Steele); Centre for Clinical Epidemiology, Lady Davis Institute for Medical Research, Jewish General Hospital, McGill University, Montreal, Quebec, Canada (Ian Shrier); Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University, Montreal, Quebec, Canada (Jay S. Kaufman, Robert W. Platt); Department of Pediatrics, Faculty of Medicine, McGill University, Montreal, Quebec, Canada (Robert W. Platt); and Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada (Robert W. Platt).

This study did not receive any dedicated funding, but the authors report the following sources of support: the Lady Davis Institute for Medical Research, Jewish General Hospital (I.S.); the Canada Research Chairs Program (J.S.K.); the National Science and Engineering Research Council of Canada (R.J.S.); the Fonds de recherche du Québec–Santé (Chercheur-National Award); and the Research Institute of the MUHC, McGill University Health Centre (R.W.P.).

We thank Drs. Luciana A. C. Machado, Chris G. Maher, Rob D. Herbert, Helen Clare, and James H. McAuley for providing access to the data from their study (25).

Conflict of interest: none declared.


