

## Practice of Epidemiology

# A Simple Method for Principal Strata Effects When the Outcome Has Been Truncated Due to Death

Yasutaka Chiba\* and Tyler J. VanderWeele

\* Correspondence to Dr. Yasutaka Chiba, Department of Environmental Medicine and Behavioral Science, Kinki University School of Medicine, 377-2, Ohno-higashi, Osakasayama, Osaka 589-8511, Japan (e-mail: chibay@med.kindai.ac.jp).

*Initially submitted July 30, 2010; accepted for publication October 25, 2010.*

In randomized trials with follow-up, outcomes such as quality of life may be undefined for individuals who die before the follow-up is complete. In such settings, restricting analysis to those who survive can give rise to biased outcome comparisons. An alternative approach is to consider the “principal strata effect” or “survivor average causal effect” (SACE), defined as the effect of treatment on the outcome among the subpopulation that would have survived under either treatment arm. The authors describe a very simple technique that can be used to assess the SACE. They give both a sensitivity analysis technique and conditions under which a crude comparison provides a conservative estimate of the SACE. The method is illustrated using data from the ARDSnet (Acute Respiratory Distress Syndrome Network) clinical trial comparing low-volume ventilation and traditional ventilation methods for individuals with acute respiratory distress syndrome.

causal inference; randomized trials; stratification; truncation

Abbreviations: QOL, quality of life; SACE, survivor average causal effect.

In a number of randomized trials in which the outcome requires considerable follow-up study, participants may die before the trial is complete and the outcome is assessed. In such cases, for the individuals who die before follow-up is complete, the outcome is not simply missing but undefined. If, for example, the outcome were quality of life (QOL) at 18 months' follow-up, QOL for individuals who have died is undefined. Some authors refer to this situation as one in which the outcome is “truncated by death” (1–4) or “censored by death” (5) to distinguish this scenario from cases in which the outcome is merely missing because of inadequate data collection.

In these settings, a crude comparison of the outcome, such as QOL, between those who survived in each treatment arm may give misleading results; by conditioning on a post-treatment event, namely, survival, we no longer preserve randomization. The treatment may, for example, render survival more likely in addition to affecting QOL. A crude comparison of QOL outcomes in both groups might erroneously lead to the conclusion that the untreated individuals have a higher QOL, simply because the unhealthy individ-

uals die under the untreated condition. A treatment comparison that makes sense in this setting would be to ask how QOL differs between treated and untreated individuals in the subpopulation that would have survived under either arm. This effect is sometimes referred to as a survivor average causal effect (SACE) (6) or a principal strata effect, and this approach to handling the “truncation by death” problem is sometimes referred to as the principal stratification approach (1, 7). Unfortunately, this subpopulation of interest is not identified; if we know that an individual survived under one treatment arm, we do not know whether he or she would have survived under the other one. A variety of statistical and sensitivity analysis techniques have been developed in the causal inference literature in statistics to attempt to address this problem (1–16). Unfortunately, many of these techniques are difficult to implement in practice or require special statistical programming.

The aim of this paper is 2-fold. First, we hope to bring these concepts of principal stratification to the epidemiology literature. Second, we describe a method for the SACE that is particularly simple to use and does not require special

statistical programming. We build on related work (17) for “principal strata direct effects” in the context of mediation (18–24).

## DEFINITIONS AND NOTATION

Suppose that  $A$  denotes the treatment variable in a randomized trial that is binary, for example,  $A = 1$  indicates treated and  $A = 0$  indicates the control condition. Let  $Y$  be an outcome of interest that is measured after some follow-up period. Let  $S$  be an indicator of whether the individual is alive when the outcome  $Y$  is measured, with  $S = 1$  indicating alive and  $S = 0$  indicating dead. For individuals who died ( $S = 0$ ), the outcome  $Y$  is undefined.

For each individual, we can also consider “counterfactual outcomes” or “potential outcomes” (25, 26) corresponding to what would have happened had an individual been in the treatment arm other than the one to which he or she was assigned. Let  $S_1$  and  $S_0$  denote the survival status for each individual under treatment  $A = 1$  and  $A = 0$ , respectively. In actuality, we observe only one of  $S_1$  or  $S_0$ ; we observe  $S_1$  for an individual who in fact had  $A = 1$ , and we observe  $S_0$  for an individual who in fact had  $A = 0$ . We have no way of observing the other potential outcome. However, at least hypothetically, we can conceive of what the survival status would have been for each individual under each of the 2 possible treatment scenarios. Likewise, we let  $Y_1$  and  $Y_0$  denote the outcome  $Y$  for each individual under treatment  $A = 1$  and  $A = 0$ , respectively. The variables  $Y_1$  and  $Y_0$  are defined only if  $S_1 = 1$  and  $S_0 = 1$ , respectively. Otherwise, the individual would have died and the outcome  $Y$  would be undefined.

A crude comparison of the outcome  $Y$  among the treated and the controls would consist of, for example, comparing the means of  $Y$  in each treatment arm among those who in fact survived:

$$E[Y|A = 1, S = 1] - E[Y|A = 0, S = 1].$$

Note that  $E[Y|A = 1, S = 1]$  is estimated by the sample mean of  $Y$  among those surviving in the treated group, and  $E[Y|A = 0, S = 1]$  is estimated by the sample mean of  $Y$  among those surviving in the control arm.

As noted at the beginning of this paper, the simple crude contrast given above is not a fair comparison because the group that survived without treatment may be healthier overall than those who survived with treatment. The control condition may have resulted in unhealthy individuals dying but for whom treatment would have kept alive. These less healthy individuals who would have died under the control condition but survived under treatment are included in the average outcomes (e.g., QOL) when examining the treated individuals who survived but would not be included in the average when examining the controls who survived. The crude comparison above effectively compares outcomes for different populations, not for the same population comparing different treatments. A simple example demonstrating why a different approach is needed is given in Appendix 1.

As an alternative to the crude measure, one can assess the “principal strata effect” or SACE. This is defined below as in prior literature (1, 7, 8).

**Definition 1:** The principal strata effect or SACE is defined as the effect of treatment among the subpopulation that would have survived under either treatment arm:

$$\text{SACE} = E[Y_1 - Y_0 | S_1 = S_0 = 1].$$

The SACE compares the outcome  $Y$  under the treated versus the control condition but among only the subpopulation that would have survived irrespective of which treatment arm they were assigned. A subpopulation such as this one that is defined by reference to potential outcomes under 2 different treatment scenarios is referred to as a “principal stratum” (7).

By restricting the comparison to the subpopulation that would have survived under either treatment arm, we circumvent the problem with the crude comparison that, for the treatment group, we include potentially less healthy individuals who would have died if they had been in the control arm. Trying to identify and estimate the SACE from data is subject to the challenge that we do not know which individuals would have survived under either treatment arm. In the next section, we describe a very simple method that can be used to try to assess the SACE. The analysis is facilitated by what is sometimes referred to as a “monotonicity” assumption:

**Assumption 1 (monotonicity):** For all individuals,  $S_0 \leq S_1$ .

Assumption 1 states that, for all individuals, survival under the treatment condition is always at least as good as survival under the control condition. In other words, survival under control implies survival under treatment, and death under treatment implies death under control. If treatment cannot render death more likely than the control condition for any individual, this assumption will be reasonable. Note that the assumption would also hold if treatment had no effect on survival. In the following section, we assume that this assumption holds. In Appendix 2, we describe a method that can be used even when this monotonicity assumption does not hold.

## A SIMPLE METHOD FOR THE SACE

Our main result in this paper expresses the SACE as the difference between the crude comparison of the outcomes across treatment arms among survivors and a sensitivity analysis parameter. We state the result and then describe its interpretation and use. A proof of this result is given in Web supplement 1. (This information is described in the first of 2 online supplements; each is referred to as “Web supplement” in the text and is posted on the *Journal's* Web site (<http://aje.oupjournals.org/>).

**Result 1:** Suppose that treatment  $A$  is randomized and that the monotonicity assumption (assumption 1) holds; then,

$$\text{SACE} = E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1] - \alpha,$$

where  $\alpha = E[Y_1 | A = 1, S = 1] - E[Y_1 | A = 0, S = 1]$ .

The result states that, to obtain the SACE, one can use the crude difference in outcomes  $Y$  between the treated and control conditions among those who survived,  $E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1]$ , and then subtract the sensitivity analysis parameter  $\alpha$ . The sensitivity analysis parameter  $\alpha$  is set by the investigator according to what is thought plausible. The sensitivity analysis parameter can be varied over a range of plausible values to examine how conclusions vary under different values for the parameter. Note that this result holds only under the monotonicity assumption (assumption 1). Otherwise, the SACE is not simply the difference between the crude estimate and the sensitivity analysis parameter  $\alpha$ , and a more complex sensitivity analysis is required. We describe such an approach under violations of the monotonicity assumption in Appendix 2.

To obtain the confidence interval for the SACE for a fixed value of parameter  $\alpha$ , one can simply subtract  $\alpha$  from the upper and lower confidence limits for  $E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1]$ .

The parameter itself is the average difference in the outcome that would have been observed under treatment comparing 2 different populations: the first is the population that would have survived with treatment ( $A = 1, S = 1$ ); the second is the population that would have survived without treatment ( $A = 0, S = 1$ ). Because the second population consists of individuals who survived even without treatment, it will likely overall be healthier than the population that would have survived with treatment. The interpretation of  $\alpha$  then is simply the difference in expected outcomes under treatment for these 2 populations.

The fact that the population that would have survived without treatment is likely healthier overall than the population that would have survived with treatment will help us derive a second result. This second result will essentially show that, in certain circumstances, the crude comparison of the outcome is conservative for the SACE. We will need one further assumption:

Assumption 2:  $E[Y_1 | A = 1, S = 1] - E[Y_1 | A = 0, S = 1] \leq 0$ .

This second assumption, which will be used in our second result below, requires that the sensitivity analysis parameter  $\alpha = E[Y_1 | A = 1, S = 1] - E[Y_1 | A = 0, S = 1]$  be less than or equal to 0. If the outcome  $Y$  were QOL and if it were indeed the case that the population that would have survived without treatment ( $A = 0, S = 1$ ) is healthier overall than the population that would have survived with treatment ( $A = 1, S = 1$ ), and if owing to the fact that this former group was healthier it also would have had higher QOL outcomes under treatment, then assumption 2 will be satisfied. Under assumptions 1 and 2, the crude comparison of the outcomes  $Y$  between the treated and control conditions among those who survived will give a lower bound for the SACE.

Result 2: Suppose that treatment  $A$  is randomized and that assumptions 1 and 2 hold; then,

**Table 1.** Data From the ARDSnet Clinical Trial<sup>a</sup>

	Low-Volume Group		Traditional-Ventilation Group	
	No.	%	No.	%
Patients	473		429	
180-Day survivors	327		256	
Proportion of survivors		69		60
Mean days to return home (survivors)	33.55		40.70	

<sup>a</sup> Refer to reference 27 for further information about this trial.

$$\text{SACE} \geq E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1].$$

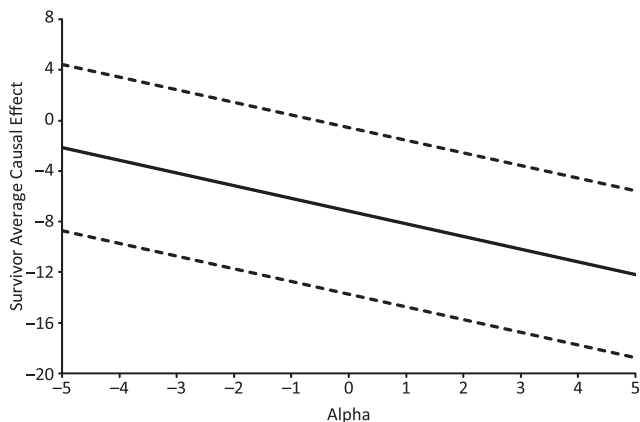
Result 2 follows immediately from result 1 and assumption 2. If QOL were the outcome of interest, then, under assumptions 1 and 2, we could use the data to estimate  $E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1]$ . Furthermore, we would know from result 2 that this crude estimate was conservative for the SACE, that is, conservative for the extent to which the treatment increased QOL among the subpopulation that would have survived irrespective of whether or not treatment was given.

Note that, if assumption 2 is reversed so that  $E[Y_1 | A = 1, S = 1] - E[Y_1 | A = 0, S = 1] \geq 0$ , then the conclusion of result 2 would be modified to  $\text{SACE} \leq E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1]$ . If the outcome is QOL, then result 2 as stated above will be the result of interest. In the following section, we illustrate where the reverse of assumption 2 and the reverse conclusion of result 2 is the setting that is applicable.

## ILLUSTRATION

The ARDSnet clinical trial (27) compared 2 methods of ventilation for patients with acute lung injury and acute respiratory distress syndrome. The 2 methods compared were low-volume ventilation ( $A = 1$ ) and traditional-volume ventilation ( $A = 0$ ). Table 1 describes the ARDSnet patients and the outcomes between the treatment arms. The study found a significant decrease in 180-day mortality ( $P = 0.003$ ) comparing the low-volume group (146/473; 31%) with the traditional-ventilation group (173/429; 40%). Note that the ARDSnet clinical trial was controversial because some commentators in the literature thought that the “standard of care” for high tidal volume was set unethically high so as to assure a positive study finding (28).

As shown in Table 1, the study also assessed a variety of outcomes that were defined for only those who had survived up through 180 days. One of these outcomes was number of days to return home. It was found that those in the low-volume group required fewer days ( $-7.15$ ;  $P = 0.03$ ; 95% confidence interval:  $-13.73, -0.56$ ) to return home (mean: 33.55) compared with the traditional-ventilation group (mean: 40.70). These means were calculated among those



**Figure 1.** Sensitivity analysis of the survivor average causal effect in the ARDSnet clinical trial (27). Solid line: estimated survivor average causal effect; broken lines: 95% confidence interval.

who survived 180 days. Because the proportion of survival in the low-volume group was higher, some of the individuals who survived in the low-volume group may have died had they been in the traditional-ventilation group. The crude comparisons of means may thus not be an adequate measure of the extent to which the low-volume-ventilation method decreases days to return home. We might thus instead be interested in the effect of low-volume ventilation versus traditional ventilation among the subset that would have survived under either ventilation method, that is, the SACE.

The low-volume ventilation method significantly reduced mortality, suggesting that the monotonicity assumption (assumption 1) may be reasonable; we will, however, return to this point below. Under assumption 1, we can apply result 1 to yield estimates of the SACE under different specifications of the sensitivity analysis parameter:

$$\alpha = E[Y_1 | A = 1, S = 1] - E[Y_1 | A = 0, S = 1].$$

The sensitivity analysis parameter compares the days to return home under the low-volume-ventilation method between 2 populations: the population that would have survived under low-volume ventilation and the population that would have survived under traditional ventilation. Because traditional ventilation is more likely to result in mortality, those who would have survived under traditional ventilation are likely a healthier population and one more likely to return home sooner if given low-volume ventilation. If we thought that the difference between these populations were small, we might specify a difference to return home of  $\alpha = 1$  day, which, by result 1, would give an estimate of the SACE of  $-8.15$  (95% confidence interval:  $-14.73, -1.56$ ); if we thought that the difference in the populations were somewhat larger, we might specify a difference of  $\alpha = 4$  days, which, by result 1, would give an estimate of the SACE of  $-11.15$  (95% confidence interval:  $-17.73, -4.56$ ). Figure 1 depicts how estimates of the SACE change as the sensitivity analysis parameter  $\alpha$  changes.

Irrespective of the actual value of  $\alpha$ , if we thought that the population that would survive under traditional ventilation was indeed healthier and would return home sooner with low-volume ventilation than the population that would survive under low-volume ventilation, then we would have that  $\alpha \geq 0$ , that is, the reverse of assumption 2. We thus would have the reverse conclusion of result 2 that the SACE was in fact less than or equal to the crude estimate,  $E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1]$ . We would then conclude that  $-7.15$  days ( $P = 0.03$ ; 95% confidence interval:  $-13.73, -0.56$ ) was an upper bound on, that is, conservative for, the SACE.

The data from the ARDSnet clinical trial were also analyzed within a principal stratification framework by Hayden et al. (6). However, their technique was considerably more difficult to implement, and it required specification of 2 cumulative proportional odds models and use of a minimization algorithm. The method used by Hayden et al. did not impose the monotonicity assumption. Although the significant reduction in mortality among the low-volume group suggests that monotonicity may be reasonable here, this assumption cannot be verified from data. In Appendix 2, we describe a relatively simple approach to assess the SACE that does not require the monotonicity assumption.

## DISCUSSION

In this paper, we have described a simple technique that can be used to assess the SACE. Analysis of such an effect is important when the outcome is potentially “truncated” or undefined for individuals who die before the outcome occurs or is measured. Under a monotonicity assumption, our method requires only that the investigator specify a single sensitivity analysis parameter. Under certain assumptions, our method also gives the result that the crude estimator is conservative for the SACE.

Using the additive scale for the sensitivity analysis parameter, as in result 1, has 2 advantages: 1) the formula for the SACE is very simple, and thus our method is very easy to implement in practice; and 2) once the sensitivity parameter has been fixed, the standard error is the same as that of the crude estimate, and thus confidence intervals are obtained immediately by just subtracting  $\alpha$  from both limits of the confidence interval for the crude estimate. By using the other parameterizations such as multiplicative sensitivity analysis parameters, one may be able to obtain a simple SACE formula, but obtaining the confidence intervals is generally more difficult than what is required by our additive parameterization.

Throughout this paper, we have focused on the setting of a clinical trial in which the primary treatment of interest is randomized. Our results also hold in an observational study, if the effect of the treatment on the outcome is unconfounded conditional on some set of covariates  $C$ . If the effect of the treatment on the outcome is unconfounded conditional on  $C$ , then results 1 and 2 hold conditional on  $C$  and assumption 2 also needs to be modified to be conditional on  $C$ . However, these results carry over in a very straightforward way.

Analysis of the SACE (or principal strata effect) is important in assessing QOL outcomes in settings in which some of the participants may die. Similarly, analysis of the SACE is important in assessing the cost of different treatment options when individuals may die before the full costs of different treatments are incurred. If one pursues a naive analysis restricted to individuals who survive, there may be situations in which treatment *A* is less expensive than treatment *B* for every individual who would have survived under both treatments, but it appears that treatment *B* is less expensive in the crude comparison: Individuals who are unhealthy and potentially quite costly might die under treatment *B* before the cost is incurred. In such settings, it may be informative to present both the crude and the principal strata analysis because the former may still be of use in cost-effectiveness analyses concerning dollars per life-year saved.

Principal strata effects have also been of use in infectious disease contexts and vaccine trials in which an outcome (e.g., human immunodeficiency virus viral load) is defined only for those individuals who, during the trial, are infected (9–12). A number of further applications concerning principal strata effects have also been pursued in the literature (4–6, 13–16). We hope that the contributions in this paper will help bring these ideas to epidemiologists and provide them with a simple tool for analyzing these principal strata effects.

## ACKNOWLEDGMENTS

Author affiliations: Department of Environmental Medicine and Behavioral Science, Kinki University School of Medicine, Osaka, Japan (Yasutaka Chiba); Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts (Tyler J. VanderWeele); and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts (Tyler J. VanderWeele).

This work was supported by National Institutes of Health grant HD060696.

Conflict of interest: none declared.

## REFERENCES

- Zhang JL, Rubin DB. Estimation of causal effects via principal stratification when some outcomes are truncated by “death.” *J Educ Behav Stat.* 2003;28(4):353–368.
- Imai K. Sharp bounds on causal effects in randomized experiments with “truncation-by-death.” *Stat Probabil Lett.* 2008;78(2):144–149.
- Kurland BF, Heagerty PJ. Directly parameterized, regression conditioning on being alive: analysis of longitudinal data truncated by deaths. *Biostatistics.* 2005;6(2):241–258.
- Mattei A, Mealli F. Application of the principal stratification approach to the Faenza randomized experiment on breast self-examination. *Biometrics.* 2007;63(2):437–446.
- Rubin DB. Causal inference through potential outcomes and principal stratification: application to studies with “censoring” due to death (with discussion). *Stat Sci.* 2006;21(3):299–321.
- Hayden D, Pauler DK, Schoenfeld D. An estimator for treatment comparisons among survivors in randomized trials. *Biometrics.* 2005;61(1):305–310.
- Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics.* 2002;58(1):21–29.
- Robins JM. A new approach to causal inference in mortality studies with sustained exposure period—application to control of the healthy worker survivor effect. *Math Model.* 1986;7:1393–1512.
- Gilbert PB, Bosch RJ, Hudgens MG. Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials. *Biometrics.* 2003;59(3):531–541.
- Hudgens MG, Hoering A, Self SG. On the analysis of viral load endpoints in HIV vaccine trials. *Stat Med.* 2003;22(14):2281–2298.
- Hudgens MG, Halloran ME. Causal vaccine effects on binary postinfection outcomes. *J Am Stat Assoc.* 2006;101(473):51–64.
- Shepherd BE, Gilbert PB, Jemai Y, et al. Sensitivity analyses comparing outcomes only existing in a subset selected post-randomization, conditional on covariates, with application to HIV vaccine trials. *Biometrics.* 2006;62(2):332–342.
- Shepherd BE, Gilbert PB, Lumley T. Sensitivity analyses comparing time-to-event outcomes existing only in a subset selected post-randomization. *J Am Stat Assoc.* 2007;102(478):573–582.
- Jemai Y, Rotnitzky A, Shepherd BE, et al. Semiparametric estimation of treatment effects given base-line covariates on an outcome measured after a post-randomization event occurs. *J R Stat Soc Series B Stat Methodol.* 2007;69(5):879–901.
- Egleston BL, Scharfstein DO, Freeman EE, et al. Causal inference for non-mortality outcomes in the presence of death. *Biostatistics.* 2007;8(3):526–545.
- Frangakis CE, Rubin DB, An MW, et al. Principal stratification designs to estimate input data missing due to death. *Biometrics.* 2007;63(3):641–649; discussion 650–662.
- Chiba Y. Bias analysis for the principal stratum direct effect in the presence of confounded intermediate variables. *J Biomet Biostat.* 2010;1(1):101.
- Rubin DB. Direct and indirect effects via potential outcomes. *Scand J Stat.* 2004;31:161–170.
- VanderWeele TJ. Simple relations between principal stratification and direct and indirect effects. *Stat Probab Lett.* 2008;78(17):2957–2962.
- Gallop R, Small DS, Lin JY, et al. Mediation analysis with principal stratification. *Stat Med.* 2009;28(7):1108–1130.
- Sjölander A, Humphreys K, Vansteelandt S, et al. Sensitivity analysis for principal stratum direct effects, with an application to a study of physical activity and coronary heart disease. *Biometrics.* 2009;65(2):514–520.
- VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology.* 2010;21(4):540–551.
- Chiba Y. Estimating the principal stratum direct effect when the total effects are consistent between two standard populations. *Stat Probab Lett.* 2010;80(11–12):958–961.
- Robins JM, Richardson TS, Spirtes P. On identification and inference for direct effects. *Epidemiology.* In press.
- Neyman J. On the application of probability theory to agricultural experiments [1923; in French]. Excerpts reprinted in

- English (Dabrowska D, Speed T, translators). *Stat Sci*. 1990;5(4):463–472.
26. Rubin D. Estimating causal effects of treatments in randomized and non-randomized studies. *J Educ Psychol*. 1974; 66(5):688–701.
  27. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network. *N Engl J Med*. 2000;342(18):1301–1308.
  28. Mann H. Controversial choice of a control intervention in a trial of ventilator therapy in ARDS: standard of care arguments in a randomised controlled trial. *J Med Ethics*. 2005;31(9):548–553.

## APPENDIX 1

### A Hypothetical Example

Consider a hypothetical randomized trial for evaluating the effect of a treatment on QOL at 18 months' follow-up. For simplicity, the outcome is dichotomized into high and low QOL scores. Two hundred participants were randomized to a treatment group and a placebo group. As shown in Appendix Table 1, of 100 participants assigned to the treatment group, 20 died before follow-up was complete, and 40 of 80 survivors had high QOL scores. Regarding the placebo group, 50 of 100 participants died before follow-up was complete, and 10 of 50 survivors had high QOL scores. A crude comparison of the proportion of participants who had high QOL scores between 2 groups is

$$\begin{aligned} E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1] \\ = 40/80 - 10/50 = 0.3. \end{aligned}$$

This comparison would not be fair, as noted in the main text of this paper. To make a fair comparison, a principal stratification approach (1, 7) could be used. This approach considers 4 types of participants that define 4 "principal strata": 1) always survivors, who would survive irrespective of the assigned group, that is,  $S_1 = S_0 = 1$ ; 2) never survivors, who would die irrespective of the assigned group, that is,  $S_1 = S_0 = 0$ ; 3) compliers, who would survive if they were assigned to the treatment group and would die if they were assigned to the placebo group, that is,  $S_1 = 1$  and  $S_0 = 0$ ; and 4) defiers, who would die if they were assigned to the treatment group and would survive if they were assigned to the placebo group, that is,  $S_1 = 0$  and  $S_0 = 1$ . In this example, if no defiers exist, the number of always survivors, never survivors, and compliers might be as shown in Appendix Table 2.

Comparisons of QOL scores for each of these 3 principal strata are fair, because the comparisons are made between the different 2 treatment arms for the same populations. Of these 3 principal strata, we can compare QOL scores among only always survivors; for never survivors and compliers assigned to the placebo group, no survivor exists and their QOL scores cannot be defined. This comparison among always survivors with  $S_1 = S_0 = 1$  is the SACE of definition 1:  $SACE = E[Y_1 - Y_0 | S_1 = S_0 = 1]$ . Again, this comparison is

**Appendix Table 1.** Data From a Hypothetical Randomized Trial

	Treatment Group	Placebo Group
No. overall	100	100
No. of survivors	80	50
No. of survivors with high QOL	40	10

Abbreviation: QOL, quality of life.

fair because it is made between the different 2 treatment arms for the same populations. The data in Appendix Table 2 show that the estimate of the SACE is

$$SACE = E[Y_1 - Y_0 | S_1 = S_0 = 1] = 35/50 - 10/50 = 0.5.$$

Note that, unfortunately, we cannot know from the observed data which participants are always survivors in the treatment group. The situation is more complex if defiers exist and makes assessment of the SACE more difficult. In Appendix 2, however, we also provide a sensitivity analysis technique that can be used even if there are defiers.

## APPENDIX 2

### An Approach Without Monotonicity

In this Appendix, we describe a method that can be used even when the monotonicity assumption (assumption 1) does not hold, that is, when there might be individuals who would survive under the control condition but not under treatment. Our result expresses the SACE using the observed difference between the outcomes across arms among survivors but, unlike result 1, uses 3 sensitivity analysis parameters rather than just one. The result is as follows:

Result 3. Suppose that treatment *A* is randomized; then, the SACE is given by

$$\begin{aligned} SACE = E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1] \\ + \frac{\pi_{01}}{p_0} \beta_0 - \frac{p_1 - p_0 + \pi_{01}}{p_1} \beta_1, \end{aligned}$$

where  $p_a = \Pr(S = 1 | A = a)$  and where  $\beta_0$ ,  $\beta_1$ , and  $\pi_{01}$  are sensitivity analysis parameters:

$$\beta_0 = E[Y_0 | S_1 = 0, S_0 = 1] - E[Y_0 | S_1 = S_0 = 1],$$

$$\beta_1 = E[Y_1 | S_1 = 1, S_0 = 0] - E[Y_1 | S_1 = S_0 = 1],$$

and  $\pi_{01} = \Pr(S_1 = 0, S_0 = 1)$ .

A proof of this result is given in Web supplement 1. Here, we describe the interpretation and use of result 3. Result 3 states that, to obtain the SACE, one can use the crude difference in outcomes *Y* between the treated and control subjects among those who survived,  $E[Y | A = 1, S = 1] - E[Y | A = 0, S = 1]$ . The sensitivity analysis parameters  $\beta_0$ ,  $\beta_1$ , and  $\pi_{01}$  are set by the investigator according to what is thought plausible. The probabilities  $p_1 = \Pr(S = 1 | A = 1)$  and  $p_0 = \Pr(S = 1 | A = 0)$  can be estimated from the data. The

**Appendix Table 2.** Data From a Hypothetical Randomized Trial Under the Principal Stratification

	Treatment Group			Placebo Group		
	Always Survivor	Never Survivor	Complier	Always Survivor	Never Survivor	Complier
No. overall	50	20	30	50	20	30
No. of survivors	50	0	30	50	0	0
No. of survivors with high QOL	35	Undefined	5	10	Undefined	Undefined

Abbreviation: QOL, quality of life.

sensitivity analysis parameters can be varied over a range of plausible values to examine how conclusions vary under different values for the parameter.

The parameters  $\beta_0$  and  $\beta_1$  are the differences in the outcome that would have been observed under different treatments comparing 2 different populations. The first parameter,  $\beta_0 = E[Y_0 | S_1 = 0, S_0 = 1] - E[Y_0 | S_1 = S_0 = 1]$ , contrasts the average outcomes under the control condition between 1) the population that would have survived under control but not under treatment and 2) the population that would have survived under both treatment and control. The second parameter,  $\beta_1 = E[Y_1 | S_1 = 1, S_0 = 0] - E[Y_1 | S_1 = S_0 = 1]$ , contrasts the average outcomes under treatment between 1) the population that would have survived under treatment but not under control and 2) the population that would have survived under both population treatment and control. Note that, because the second population in both compar-

isons constitutes the group that would have survived irrespective of treatment, it is likely that, overall, this population will be healthier than the other population in each comparison, suggesting that the parameters  $\beta_0$  and  $\beta_1$  will both likely be negative. The third sensitivity analysis parameter,  $\pi_{01} = \Pr(S_1 = 0, S_0 = 1)$ , is the proportion of the defiers who would have survived under the control condition but not under treatment. The parameter  $\pi_{01}$  is simply the proportion of defiers. This proportion is assumed to be 0 by the monotonicity assumption (assumption 1) and result 1 but is allowed to be nonzero by result 3.

By varying the 3 parameters,  $\beta_0$ ,  $\beta_1$ , and  $\pi_{01}$ , a researcher can use result 3 to assess what the SACE might be, even when the monotonicity assumption (assumption 1) is violated. Some further remarks about bounds in the context of violations of the monotonicity assumption are given in Web supplement 2.