STATISTICS IN MEDICINE
*Statist. Med.* 2007; **26**:5125–5135
Published online 24 May 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.2927

# Bounds on potential risks and causal risk differences under assumptions about confounding parameters

Yasutaka Chiba[1,*,†], Tosiya Sato[1] and Sander Greenland[2]

[1]*Department of Biostatistics, Kyoto University School of Public Health, Kyoto, Japan*
[2]*Department of Epidemiology and Department of Statistics, University of California, Los Angeles, CA, U.S.A.*

SUMMARY

Nonparametric bounds on causal effects in observational studies are available under deterministic potential-outcome models. We derive narrower bounds by adding assumptions regarding bias due to confounding. This bias is defined as the difference between the expectation of potential outcomes for the exposed group and that for the unexposed group. We show that crude effect measures bound causal effects under the given assumptions. We then derive bounds for randomized studies with noncompliance, which are given by the per protocol effect. With perfect compliance in one treatment group, the direction of effect becomes identifiable under our assumptions. Although the assumptions are not themselves identifiable, they are nonetheless reasonable in some situations. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: instrumental-variables estimator; intention to treat; noncompliance; per protocol set; potential outcomes

## 1. INTRODUCTION

Confounding is a major concern in observational studies. Even in randomized studies, the factors affecting compliance can act as confounders. When unmeasured confounders are present, causal effects cannot be unbiasedly estimated without assumptions that are not identified from the study data alone. This problem arises in randomized trials as well: although randomization provides statistical control of confounding under perfect compliance and follow-up, nonrandom noncompliance can introduce confounding into analyses of treatment received ('as-treated' analysis).

Nonetheless, under deterministic causal models, bounds on causal effects can be estimated, and various nonparametric bounds have been proposed [1–5]. Under their assumptions and

*Correspondence to: Yasutaka Chiba, Department of Biostatistics, Kyoto University School of Public Health, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan.
†E-mail: chibay@pbh.med.kyoto-u.ac.jp

*Received 3 August 2006*
*Accepted 10 April 2007*

no more (which implicitly include assumptions of no selection bias and no misclassification), the bounds derived by Tian and Pearl [5] are the 'sharpest,' i.e. have the narrowest width of any such bounds. Using assumptions on measured confounding, however, MacLehose *et al.* [6] derived stratum-specific bounds narrower than the Tian–Pearl bounds in each stratum.

We extend here the previous work as follows. In Section 2, we introduce two assumptions to derive bounds narrower than those of Tian and Pearl [5], and apply them to a cohort study of personality and heart disease. In Section 3, we translate these assumptions and bounds to randomized studies with noncompliance, and apply them to a randomized trial of vitamin A and childhood mortality. Section 4 provides some discussion, including plausibility considerations.

Similar to previous authors, we derive bounds on expectations and hence leave aside issues of random error in the bounds estimated from data. Random error can be straightforwardly incorporated using normal approximations or bootstrapping; we employ the former. We use $E$ for a nonrandomized exposure indicator, and assume the now standard deterministic potential-outcome model in which $Y(1)$ and $Y(0)$ are the potential-outcome indicators under $E=1$ and 0, respectively [7, 8], a model used in several textbooks (e.g. References [9–11]). The 'potential risks' $\Pr(Y(1)=1)$ and $\Pr(Y(0)=1)$ are then the expectations of $Y$ if everyone in the total study cohort is exposed or given $E=1$, and if everyone is not exposed or given $E=0$, respectively.

Causal effects are contrasts of these expectations, with the causal risk difference or 'average causal effect' (ACE) defined as

$$\mathrm{RD}_{\mathrm{causal}} = \Pr(Y(1)=1) - \Pr(Y(0)=1)$$

Note that the observed outcome $Y$ equals the potential outcome $Y(e)$ whenever $E=e$. Hence, the crude association measure corresponding to $\mathrm{RD}_{\mathrm{causal}}$ is

$$\begin{aligned}\mathrm{RD}_{\mathrm{crude}} &= \Pr(Y=1|E=1) - \Pr(Y=1|E=0)\\ &= \Pr(Y(1)=1|E=1) - \Pr(Y(0)=1|E=0)\end{aligned}$$

The bias due to confounding is then $\mathrm{RD}_{\mathrm{crude}} - \mathrm{RD}_{\mathrm{causal}}$.

## 2. BOUNDS FOR NONRANDOMIZED COHORT STUDIES

### *2.1. Existing bounds*

The Tian–Pearl bounds on $\Pr(Y(e)=1)$ are

$$\Pr(Y=1, E=e) \leqslant \Pr(Y(e)=1) \leqslant 1 - \Pr(Y=0, E=e)$$

where $e=0, 1$ [5]. While these bounds were originally derived *via* linear programming, a simpler derivation [12] notes that $\Pr(Y(e)=y, E=e) = \Pr(Y=y, E=e)$ and hence

$$\Pr(Y(e)=1) = \Pr(Y=1, E=e) + \Pr(Y(e)=1|E=1-e)\Pr(E=1-e) \tag{1}$$

Bounds for $\mathrm{RD}_{\mathrm{causal}}$ then follow by substitution of 0 or 1 for $\Pr(Y(e)=1|E=1-e)$ into equation (1).

The width of the bounding interval for $\Pr(Y(e)=1)$ is $\Pr(E=1-e)$, with a resulting interval width for $\mathrm{RD}_{\mathrm{causal}}$ of $\Pr(E=0) + \Pr(E=1) = 1$. Since the logical limits for $\mathrm{RD}_{\mathrm{causal}}$ are $-1$



DOI: 10.1002/sim

and 1, these bounds do not even indicate the direction of effect, yet they are the narrowest (sharpest) derivable in the absence of additional assumptions [5]. Thus, Tian and Pearl [5] derived bounds under exposure-specific monotonicity assumptions (nonnegative or nonpositive). Assuming a nonnegative average effect of $E=1$, this becomes

$$\Pr(Y(1)=1|E=e)\geqslant\Pr(Y(0)=1|E=e)$$

Nonpositive average effects can be addressed by reversing the inequality. The nonnegative assumption yields the improved bounds $\Pr(Y(1)=1)\geqslant\Pr(Y=1)$ and $\Pr(Y(0)=1)\leqslant\Pr(Y=1)$, with $\mathrm{RD}_{\mathrm{causal}}\geqslant 0$; these results are reversed assuming nonpositivity.

### 2.2. *Proposed bounds*

To derive bounds that are narrower than those of Tian and Pearl [5], we introduce assumptions about confounding.

*2.2.1. Biases due to confounding.* Define the bias factors $\alpha$ and $\beta$ by

$$\alpha\equiv\Pr(Y(1)=1|E=1)-\Pr(Y(1)=1|E=0)$$

$$\beta\equiv\Pr(Y(0)=1|E=1)-\Pr(Y(0)=1|E=0)$$

$\alpha$ is the difference between the expected outcome of those actually exposed and the expected outcome of the unexposed if they were exposed. Similarly, $\beta$ is the difference between the expected outcome of exposed under nonexposure and the expected outcome of those actually unexposed.

Let $Z$ be the vector of all unmeasured confounders or a sufficient subset [13]. Robins [14] and Brumback *et al.* [15] suggested the expression

$$E[Y(e)|E=e,Z=z]-E[Y(e)|E=1-e,Z=z]$$

as a measure of confounding. $\alpha$ and $\beta$ follow for an indicator $Y$ upon averaging (standardizing) over $\Pr(Z=z)$. When $\alpha>0$ and $\beta>0$,

$$\Pr(Y(e)=1|E=1)>\Pr(Y(e)=1|E=0)$$

which means that the subjects in the exposed group tend to be sicker than those in the unexposed group (and which leads to positive confounding). When $\alpha<0$ and $\beta<0$,

$$\Pr(Y(e)=1|E=1)<\Pr(Y(e)=1|E=0)$$

which means that the subjects in the unexposed group tend to be sicker than those in the exposed group (and which leads to negative confounding). No confounding occurs when $\alpha=\beta=0$.

Bounds on $\alpha$ and $\beta$ are

$$-\Pr(Y=0|E=1)\leqslant\alpha\leqslant\Pr(Y=1|E=1) \tag{2}$$

$$-\Pr(Y=1|E=0)\leqslant\beta\leqslant\Pr(Y=0|E=0) \tag{3}$$



DOI: 10.1002/sim

because $0 \leqslant \Pr(Y(e) = 1 | E = 1 - e) \leqslant 1$. Substitution of $\alpha$ and $\beta$ into equation (1) yields

$$\Pr(Y(1) = 1) = \Pr(Y = 1 | E = 1) - \alpha \Pr(E = 0) \tag{4}$$

$$\Pr(Y(0) = 1) = \Pr(Y = 1 | E = 0) + \beta \Pr(E = 1) \tag{5}$$

Substitution of equations (2) and (3) into these equations yields the Tian–Pearl bounds.

*2.2.2. Proposed assumptions and bounds.* Equations (2)–(5) suggest ways to get sharper bounds on causal effects by introducing bounds on $\alpha$ and $\beta$. As an intuitively obvious example, consider the 'nonnegative confounding' assumption $\alpha \geqslant 0$ and $\beta \geqslant 0$, which says that potential risks for those actually exposed are no less than for those actually unexposed, i.e.

$$\Pr(Y(e) | E = 1) \geqslant \Pr(Y(e) | E = 0) \quad \text{for } e = 0, 1$$

Upon substitution into equations (4) and (5), we obtain

$$\Pr(Y(1) = 1) \leqslant \Pr(Y = 1 | E = 1)$$

and

$$\Pr(Y(0) = 1) \geqslant \Pr(Y = 1 | E = 0)$$

Then, combinations of these bounds and the Tian–Pearl bounds yield the bounding-interval width for $\Pr(Y(e) = 1)$ of $\Pr(Y = a | E = a) \Pr(E = 1 - a)$. It follows that $\mathrm{RD}_{\mathrm{causal}} \leqslant \mathrm{RD}_{\mathrm{crude}}$ when $\alpha \geqslant 0$ and $\beta \geqslant 0$.

Similarly, suppose $\alpha \leqslant 0$ and $\beta \leqslant 0$, which says that potential risks for those actually exposed are no more than for those actually unexposed, i.e.

$$\Pr(Y(e) | E = 1) \leqslant \Pr(Y(e) | E = 0) \quad \text{for } e = 0, 1$$

Then,

$$\Pr(Y(1) = 1) \geqslant \Pr(Y = 1 | E = 1)$$

$$\Pr(Y(0) = 1) \leqslant \Pr(Y = 1 | E = 0)$$

and the bounding-interval width for $\Pr(Y(e) = 1)$ is $\Pr(Y = 1 - a | E = a) \Pr(E = 1 - a)$. It follows that $\mathrm{RD}_{\mathrm{causal}} \geqslant \mathrm{RD}_{\mathrm{crude}}$ when $\alpha \leqslant 0$ and $\beta \leqslant 0$.

Now suppose $\alpha = \beta$ which says that the difference in risk between the exposed and the unexposed groups if they had been given identical exposures would not vary with exposure (identical confounding within all exposure groups). From equations (2) and (3), bounds on $\alpha = \beta$ are

$$-\min \left\{ \begin{array}{l} \Pr(Y = 0 | E = 1) \\ \\ \Pr(Y = 1 | E = 0) \end{array} \right\} \leqslant \alpha = \beta \leqslant \min \left\{ \begin{array}{l} \Pr(Y = 1 | E = 1) \\ \\ \Pr(Y = 0 | E = 0) \end{array} \right\}$$



DOI: 10.1002/sim

Substitution into equations (4) and (5) yields bounds on $\Pr(Y(e)=1)$. In particular, when $\Pr(Y=1|E=1)\leqslant 0.5$ and $\Pr(Y=1|E=0)\leqslant 0.5$, bounds on $\Pr(Y(e)=1)$ and $\mathrm{RD}_{\mathrm{causal}}$ are

$$\Pr(Y=1,\,E=1)\leqslant \Pr(Y(1)=1)\leqslant \Pr(Y=1|E=1)+\Pr(Y=1,\,E=0)$$

$$\Pr(Y=1,\,E=0)\leqslant \Pr(Y(0)=1)\leqslant \Pr(Y=1|E=0)+\Pr(Y=1,\,E=1)$$

$$-\Pr(Y=1|E=0)\leqslant \mathrm{RD}_{\mathrm{causal}}\leqslant \Pr(Y=1|E=1)$$

The bounding-interval width for $\Pr(Y(e)=1)$ is

$$\{\Pr(Y=1|E=1)+\Pr(Y=1|E=0)\}\Pr(E=1-e)$$

and that for $\mathrm{RD}_{\mathrm{causal}}$ is

$$\Pr(Y=1|E=1)+\Pr(Y=1|E=0)$$

### *2.3. An example*

Table I presents data from 22-year follow-up in a classic observational cohort study, the Western Collaborative Group Study [16], which examined the effect of behaviour patterns on coronary heart disease (CHD). Type A behaviour was characterized by enhanced aggressiveness, ambitiousness, competitive drive, and chronic sense of time urgency, and type B behaviour were subjects without such behavioural characteristics.

Table II presents estimated bounds on the potential risks and the causal risk difference under the assumptions discussed above, and various combinations. Bounds were estimated using $\hat{\Pr}(.)$, probabilities estimated from the data counts, in place of $\Pr(.)$. Standard errors were derived from the estimates of $\mathrm{var}(\hat{p}\hat{q})=\mathrm{var}(\hat{p})\mathrm{var}(\hat{q})+p^2\,\mathrm{var}(\hat{q})+q^2\,\mathrm{var}(\hat{p})$ obtained by replacing $p$ and $q$ by $\hat{p}$ and $\hat{q}$ everywhere, e.g. $\hat{p}=\hat{\Pr}(Y=1|E=e)$ and $\hat{q}=\hat{\Pr}(E=e)$ for the lower Tian–Pearl bound.

The estimated bounding-interval widths ranged from 100 per cent with no assumption, to 50.4 per cent under nonnegative average effect and 51.0 per cent under nonnegative confounding, but with little overlap. Nonetheless, the width is only 1.4 per cent under nonnegative average effect plus nonnegative confounding. Thus, when combined, assumptions can have a profound impact on the bounds. The width is not narrowed by adding the assumption of $\alpha=\beta$ because upper bounds on $\Pr(Y(e)=1)$ under $\alpha\geqslant 0$ and $\beta\geqslant 0$ are smaller than those under $\alpha=\beta$.

Table I. Data from the Western Collaborative Group Study: coronary heart disease (CHD) incidence by behaviour pattern [16].

| Groups | CHD | No CHD | Totals |
|---|---|---|---|
| Type A | 119 | 1470 | 1589 |
| Type B | 95 | 1470 | 1565 |



DOI: 10.1002/sim

Table II. Estimated bounds on the potential risks of coronary heart disease and the causal risk difference under various assumptions in the Western Collaborative Group Study.

| | $\Pr(Y(1)=1)$ | | $\Pr(Y(0)=1)$ | | $RD_{causal}$ | |
|---|---|---|---|---|---|---|
| Assumptions* | Lower | Upper | Lower | Upper | Lower | Upper |
| None | 3.77 | 53.39 | 3.01 | 53.39 | −49.62 | 50.38 |
| | (0.34) | (0.89) | (0.30) | (0.89) | (0.95) | (0.94) |
| Nonnegative average effect | 6.79 | 53.39 | 3.01 | 6.79 | 0.00 | 50.38 |
| | (0.45) | (0.89) | (0.30) | (0.45) | (0.00) | (0.94) |
| $\alpha \geqslant 0$ and $\beta \geqslant 0$ | 3.77 | 7.49 | 6.07 | 53.39 | −49.62 | 1.42 |
| | (0.34) | (0.66) | (0.60) | (0.89) | (0.95) | (0.89) |
| $\alpha = \beta$ | 3.77 | 10.50 | 3.01 | 9.84 | −6.07 | 7.49 |
| | (0.34) | (0.73) | (0.30) | (0.69) | (0.77) | (0.79) |
| $\alpha = \beta \geqslant 0$ | 3.77 | 7.49 | 6.07 | 9.84 | −6.07 | 1.42 |
| | (0.34) | (0.66) | (0.60) | (0.69) | (0.77) | (0.89) |
| Nonnegative average effect | 6.79 | 7.49 | 6.07 | 6.79 | 0.00 | 1.42 |
| $\alpha \geqslant 0$ and $\beta \geqslant 0$ | (0.45) | (0.66) | (0.60) | (0.45) | (0.00) | (0.89) |
| Nonnegative average effect | 6.79 | 7.49 | 6.07 | 6.79 | 0.00 | 1.42 |
| $\alpha = \beta \geqslant 0$ | (0.45) | (0.66) | (0.60) | (0.45) | (0.00) | (0.89) |

*Note*: Bounds in per cents with standard errors in parentheses.
*All estimators considered implicitly assume absence of biases other than confounding.

## 3. BOUNDS FOR RANDOMIZED STUDIES WITH NONCOMPLIANCE

Consider a randomized study. Let $R$ be a randomization (intent-to-treat, ITT) indicator: $R=1$ for subjects randomized to treatment, and $R=0$ for subjects randomized to control. Similarly, let $X$ indicate actual (received) treatment which under protocol violations such as noncompliance may not be randomized. We call the $\Pr(X=1-r|R=r)$ probabilities of noncompliance. Finally, paralleling Section 2, let $Y$ be the observed outcome and $Y(x)$ be the potential outcome when $X=x$.

Due to randomization, the potential outcomes are independent of treatment intent $R$, i.e. $\Pr(Y(x)=y|R=r)=\Pr(Y(x)=y)$, sometimes called 'weak ignorability of treatment assignment.' Similar to previous authors [4, 17–20], we make the instrumental assumption that the potential outcomes $Y(x)$ are not functions of treatment intent $R$ except through $X$ (i.e. that $R$ affects $Y$ only through $X$), which may be reasonable in successfully blinded studies but needs to be critically evaluated in any application.

### 3.1. Existing bounds

Robins [1] and Manski [2] gave nonparametric bounds on $\Pr(Y(x)=1)$

$$\Pr(Y=1, X=x|R=r) \leqslant \Pr(Y(x)=1) \leqslant 1-\Pr(Y=0, X=x|R=r)$$

Their width is $\Pr(X=1-r|R=r)$. The derivation of these bounds parallels that of the Tian–Pearl bounds: we have

$$\begin{aligned}\Pr(Y(x)=1) &= \Pr(Y=1, X=x|R=r) \\ &\quad + \Pr(Y(r)=1|X=1-r, R=r)\Pr(X=1-r|R=r) \end{aligned} \tag{6}$$

Bounds follow by substituting 0 or 1 for $r$ into equation (6).



DOI: 10.1002/sim

Assuming preventive average effects among both $R=1$ and $R=0$ compliers, i.e.

$$\Pr(Y(1)=1|X=R=r)\leqslant\Pr(Y(0)=1|X=R=r)$$

induces sharper bounds

$$\Pr(Y(1)=1)\leqslant\Pr(Y=1|R=1)$$

and

$$\Pr(Y(0)=1)\geqslant\Pr(Y=1|R=0)$$

[4]. The upper bound on $\mathrm{RD}_{\mathrm{causal}}$ then equals the ITT difference, i.e.

$$\mathrm{RD}_{\mathrm{causal}}\leqslant\Pr(Y=1|R=1)-\Pr(Y=1|R=0)$$

[17]. Similarly, assuming causative average effects among compliers leads to the ITT difference as the lower bound on $\mathrm{RD}_{\mathrm{causal}}$, i.e.

$$\mathrm{RD}_{\mathrm{causal}}\geqslant\Pr(Y=1|R=1)-\Pr(Y=1|R=0)$$

*3.2. Proposed bounds*

Define the $R$-specific biases due to confounding as

$$\alpha_r\equiv\Pr(Y(1)=1|X=1,\,R=r)-\Pr(Y(1)=1|X=0,\,R=r)$$

$$\beta_r\equiv\Pr(Y(0)=1|X=1,\,R=r)-\Pr(Y(0)=1|X=0,\,R=r)$$

where $r=0,1$. $\alpha_r$ and $\beta_r$ are confounding effects that would arise from within-$R$ ($R$-stratified) comparisons of those with $X=1$ *versus* those with $X=0$. The directions and magnitudes of $\alpha_r$ and $\beta_r$ may be interpreted in a manner parallel to those of $\alpha$ and $\beta$ in Section 2.2.1.

Substituting $\alpha_r$ and $\beta_r$ into equation (6) yields

$$\Pr(Y(1)=1)=\Pr(Y=1|X=1,\,R=r)-\alpha_r\Pr(X=0|R=r) \tag{7}$$

$$\Pr(Y(0)=1)=\Pr(Y=1|X=0,\,R=r)+\beta_r\Pr(X=1|R=r) \tag{8}$$

Now consider various assumptions of $R$-specific positivity of confounding: $\alpha_r\geqslant 0$ and $\beta_r\geqslant 0$. Using derivations parallel to those in Section 2.2.2, $\Pr(Y(1)=1)\leqslant\Pr(Y=1|X=1,\,R=1)$ from $\alpha_1\geqslant 0$, and $\Pr(Y(0)=1)\geqslant\Pr(Y=1|X=0,\,R=0)$ from $\beta_0\geqslant 0$. Together these yield

$$\mathrm{RD}_{\mathrm{causal}}\leqslant\Pr(Y=1|X=1,\,R=1)-\Pr(Y=1|X=0,\,R=0)$$

which compares outcomes only for compliers, the 'per protocol set' (PPS) association [17]. From positivity, the bounding-interval width for $\Pr(Y(x)=1)$ is $\Pr(Y=r|X=R=r)\Pr(X=1-r|R=r)$.



DOI: 10.1002/sim

Next, consider various assumptions of $R$-specific nonpositivity of confounding: $\alpha_r \leqslant 0$ and $\beta_r \leqslant 0$. Under $\alpha_1 \leqslant 0$ and $\beta_0 \leqslant 0$, $\Pr(Y(1)=1) \geqslant \Pr(Y=1|X=1, R=1)$ and $\Pr(Y(0)=1) \leqslant \Pr(Y=1|X=0, R=0)$. Thus,

$$\mathrm{RD}_{\mathrm{causal}} \geqslant \Pr(Y=1|X=1, R=1) - \Pr(Y=1|X=0, R=0)$$

which is the PPS measure.

Now suppose we observe perfect compliance in $R=0$ ('control') group. Then $\Pr(Y(0)=1) = \Pr(Y=1|R=0)$, and we only need to consider equation (7) to estimate effects. If we now assume $\alpha_1 \geqslant 0$, we have $\Pr(Y(1)=1) \leqslant \Pr(Y=1|X=1, R=1)$. If we assume instead $\alpha_1 \leqslant 0$, we have $\Pr(Y(1)=1) \geqslant \Pr(Y=1|X=1, R=1)$. If we make the following 'monotonic confounding' assumption:

$$\text{Either } \alpha_1 \geqslant 0 \text{ and } \beta_1 \geqslant 0 \quad \text{or} \quad \alpha_1 \leqslant 0 \text{ and } \beta_1 \leqslant 0$$

whether $\alpha_1$ is positive or negative can be identified, because we can solve for $\beta_1$ as the identified quantity

$$\frac{\Pr(Y=1|R=0) - \Pr(Y=1|X=0, R=1)}{\Pr(X=1|R=1)}$$

from equation (8).

In general, we cannot identify from the data whether

$$\Pr(Y(x)=1|X=1, R=1) \geqslant \Pr(Y(x)=1|X=0, R=1)$$

i.e. whether the subjects in $R=1$ group taking the experimental treatment tend to experience the event more frequently than the subjects in $R=1$ group taking the control treatment. We can, however, identify this inequality when we observe perfect compliance in $R=0$ group and assume monotonic confounding in the $R=1$ group, for then $\beta_1$ can be estimated. We then obtain

$$\Pr(Y(x)=1|X=1, R=1) \geqslant \Pr(Y(x)=1|X=0, R=1) \quad \text{if } \alpha_1 \geqslant 0 \quad \text{and} \quad \beta_1 \geqslant 0$$

and

$$\Pr(Y(x)=1|X=1, R=1) \leqslant \Pr(Y(x)=1|X=0, R=1) \quad \text{if } \alpha_1 \leqslant 0 \quad \text{and} \quad \beta_1 \leqslant 0$$

Now suppose we assume $\alpha_1 = \beta_1$ and $\alpha_0 = \beta_0$. This pair of equations along with equations (7) and (8) allow point solutions for $\Pr(Y(x)=1)$, with

$$\mathrm{RD}_{\mathrm{causal}} = \Pr(Y(1)=1) - \Pr(Y(0)=1) = \frac{\Pr(Y=1|R=1) - \Pr(Y=1|R=0)}{\Pr(X=1|R=1) - \Pr(X=1|R=0)}$$

the classical instrumental-variable estimator, which for binary data is usually derived by assuming instead that $R \geqslant X$ for all units, or a similar assumption [18, 19, 21].



DOI: 10.1002/sim

Table III. Data from trial of vitamin A supplementation and childhood mortality, stratified by compliance.

| | Vitamin A compliance | | Control compliance | |
|---|---|---|---|---|
| Status | Yes | No | Yes | No |
| Alive | 9663 | 2385 | 11 514 | — |
| Dead | 12 | 34 | 74 | — |
| Totals | 9675 | 2419 | 11 588 | — |

Table IV. Estimated bounds on the potential risk from active treatment and the causal risk difference, under various assumptions in the vitamin A supplementation trial.

| | $\Pr(Y(1)=1)$ | | $\mathrm{RD}_{\mathrm{causal}}$ | |
|---|---|---|---|---|
| Assumptions* | Lower | Upper | Lower | Upper |
| None | 0.10 | 20.10 | −0.54 | 19.46 |
| | (0.03) | (0.36) | (0.08) | (0.37) |
| Preventive average effect | 0.10 | 0.36 | −0.54 | −0.27 |
| | (0.03) | (0.06) | (0.08) | (0.09) |
| Monotonic confounding | 0.12 | 20.10 | −0.51 | 19.46 |
| | (0.04) | (0.36) | (0.08) | (0.37) |
| Preventive average effect | 0.12 | 0.36 | −0.51 | −0.27 |
| + monotonic confounding | (0.04) | (0.06) | (0.08) | (0.09) |
| Classical IV estimator | 0.32 | 0.32 | −0.32 | −0.32 |
| | (0.14) | (0.14) | (0.12) | (0.12) |

*Note*: Bounds in per cents with standard errors in parentheses.
*All estimators considered implicitly assume absence of biases other than confounding.

### *3.3. An example*

Table III presents data from a randomized trial of vitamin A supplementation and mortality among children from 450 Indonesian villages [22]. Although randomization was by village, following earlier illustrations we will treat the data as individually randomized, which leads to the same point estimators as cluster randomization [20]. Because the treatment ($X=1$) was not available outside the trial, there was perfect compliance among controls ($X=0$ when $R=0$).

Table IV presents estimated bounds under various assumptions. The bounds and standard errors were obtained as in Section 2 (the variance for the classical IV estimator is from Greenland [19, Appendix]). The classical IV estimator of $\mathrm{RD}_{\mathrm{causal}}$ is −0.32 per cent, which follows from assuming $\alpha_1=\beta_1$. The assumption of nonpositive effect is sufficient to bound $\mathrm{RD}_{\mathrm{causal}}$ below 0. Under the monotonic-confounding assumption, $\alpha_1$ is specified as $\alpha_1 \leqslant 0$, because $\beta_1$ is estimated as −0.96 per cent.

## 4. DISCUSSION

We have derived improved bounds on the causal risk difference by making qualitative assumptions about confounding. The bounds may be extended to bounded ordered outcomes such as



DOI: 10.1002/sim

birthweight, blood pressure, and other direct physical measurements using the general relation $\min(Y) \leqslant E[Y(x)|X = 1 - x] \leqslant \max(Y)$ in place of $0 \leqslant \Pr(Y(x) = 1|X = 1 - x) \leqslant 1$. The resulting bounds become less informative, however, if as usual $Y$ is concentrated away from its bounds.

Our results are limited insofar as the added assumptions are not identified from the data, and hence must be derived from contextual considerations. It is often the case that the direction of confounding can be argued based on external information. For example, studies of lifestyle factors and disease are expected to suffer negative (downward) confounding, because of positive correlations of healthy behaviours. Indeed, this is believed to be the problem giving rise to the repeated failure of vitamin trials to replicate negative associations seen in observational epidemiologic studies.

Confounding direction is also often anticipated in observational studies of medical interventions and pharmaceuticals. For example, when physicians prefer to reserve the new ($E = 1$) treatment for high-risk patients, positive (upward) confounding should be anticipated to the point of potentially making a helpful treatment appear harmful. Conversely, when physicians prefer to administer the new treatment to patients at low risk, negative confounding should be expected. More detailed considerations may lead to expectations of uniform confounding across subgroups. On the other hand, even if no firm expectations can be garnered from the context, the proposed bounds can be considered as part of a sensitivity analysis.

An alternative approach is to parametrically specify the structure of biases, and incorporate imprecise or uncertain external information about biases as part of a Bayesian analysis [23, 24]. The structure used here can be integrated into this format if there is sufficient information to specify priors on the $\alpha$ and $\beta$, along the lines discussed by Graham [25]. Such approaches provide direct accounting for random error along with biases in the final results.

ACKNOWLEDGEMENT

We would like to thank the referees for their helpful comments.

REFERENCES

1. Robins JM. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In *Health Service Research Methodology*: *A Focus on AIDS*, Sechrest L, Freeman H, Mulley A (eds). U.S. Public Health Service: Washington, DC, 1989; 113–159.
2. Manski CF. Nonparametric bounds on treatment effects. *American Economic Review*, *Papers and Proceedings* 1990; **80**:319–323.
3. Pearl J. Causal inference from indirect experiments. *Artificial Intelligence in Medicine* 1995; **7**:561–582.
4. Balke A, Pearl J. Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association* 1997; **92**:1171–1176.
5. Tian J, Pearl J. Probabilities of causation: bounds and identification. *Annals of Mathematics and Artificial Intelligence* 2000; **28**:287–313.
6. MacLehose RF, Kaufman S, Kaufman JS, Poole C. Bounding causal effects under uncontrolled confounding using counterfactuals. *Epidemiology* 2005; **16**:548–555.
7. Copas JB. Randomization models for matched and unmatched 2×2 tables. *Biometrika* 1973; **60**:267–276.
8. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 1974; **66**:688–701.
9. Rothman KJ, Greenland S. *Modern Epidemiology* (2nd edn). Lippincott: Philadelphia, PA, 1998.
10. Pearl J. *Causality*: *Models*, *Reasoning*, *and Inference*. Cambridge University Press: Cambridge, 2000.
11. Berk R. *Regression Analysis*: *A Constructive Critique*. Sage Publications: Thousand Oaks, CA, 2003.
12. Joffe MM. Using information on realized effects to determine prospective causal effects. *Journal of the Royal Statistical Society*, *Series B* 2001; **63**:759–774.



DOI: 10.1002/sim

13. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999; **10**:37–48.
14. Robins JM. Association, causation, and marginal structural models. *Synthese* 1999; **121**:151–179.
15. Brumback BA, Hernán MA, Haneuse SJPA, Robins JM. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Statistics in Medicine* 2004; **23**:749–767.
16. Ragland DR, Brand RJ. Coronary heart disease mortality in the Western Collaborative Group Study: follow-up experience of 22 years. *American Journal of Epidemiology* 1988; **127**:462–475.
17. Lewis B, Sheiner LB, Rubin DB. Intention-to-treat analysis and the goals of clinical trials. *Clinical Pharmacology and Therapeutics* 1995; **57**:6–15.
18. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 1986; **91**:444–455.
19. Greenland S. An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology* 2000; **29**:722–729.
20. Sommer A, Zeger SL. On estimating efficiency from clinical trials. *Statistics in Medicine* 1991; **10**:45–52.
21. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology* 2006; **17**:360–372.
22. Sommer A, Tarwotjo I, Djunaedi E, West KP, Loedin AA, Tilden R, Mele L. Impact of vitamin A supplementation on childhood mortality: a randomized controlled community trial. *Lancet* 1986; **327**:1169–1173.
23. Greenland S. The impact of prior distributions for uncontrolled confounding and response bias. *Journal of the American Statistical Association* 2003; **98**:47–54.
24. Greenland S. Multiple-bias modeling for observational studies (with Discussion). *Journal of the Royal Statistical Society*, *Series A* 2005; **168**:267–308.
25. Graham P. Bayesian inference for a generalized population attributable fraction. *Statistics in Medicine* 2000; **19**:937–956.



DOI: 10.1002/sim