Negative Controls to Detect Selection Bias and Measurement Bias in Epidemiologic Studies

Benjamin F. Arnold, Ayse Ercumen, Jade Benjamin-Chung, John M. Colford, Jr

eAPPENDIX

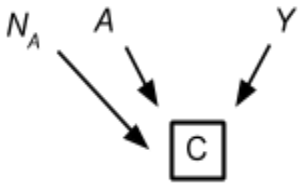
This supplement includes more detailed discussion of each example listed in the main text Table.

Example 1: Ruckart et al. Negative Control Exposure for Selection Bias

A case-control study that examined the effect of in-utero exposure to chemical contaminants in drinking water on the risk of childhood hematopoietic cancers, neural tube defects and oral clefts[1] provides an example of negative control exposures to detect the type of selection bias depicted in Figure 1A. The study launched a media campaign that encouraged families to contact investigators if they conceived a child while living in the study area during the drinking water contamination period. The investigators enrolled cases from respondents that reported one of three outcomes of interest, and randomly selected controls from the remaining respondents. Monthly average levels of water contaminants in the study area were determined using groundwater modeling, and linked to participants through their residential address. Exposure categories were defined based on average concentrations during the exposure window of interest, with the "unexposed" group defined as individuals with no residential exposure to the chemical of interest during that window. Selection bias could arise in this study if exposed cases were more likely to respond to the media campaign than exposed controls because of heightened

awareness about the health sequelae of the exposure (*A* affects *C* through increased awareness, differentially by case status *Y*). To detect this potential bias, investigators included non-relevant exposure periods as negative control exposures N_A (Figure 1A); they hypothesized that being exposed to drinking water chemicals during the third trimester of pregnancy could not plausibly cause neural tube defects and oral clefts given the timing of the formation of these organ systems. The investigators found no association between the negative control exposures and these two outcomes of interest, which helped rule out selection bias.

Figure 1A (main text) Ruckart et al.

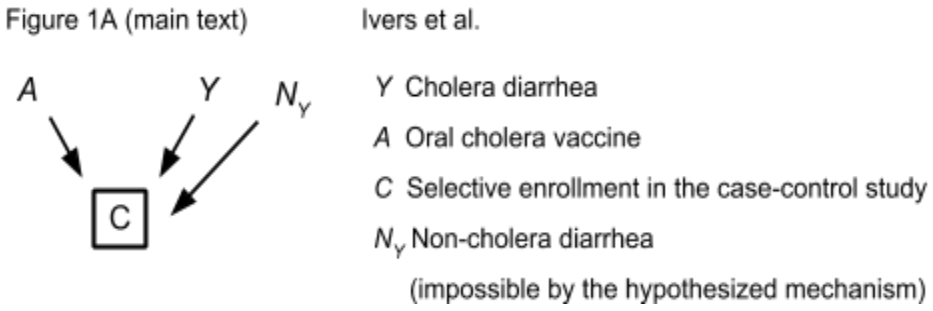


Y Hematopoietic cancer, neural tube defects, oral clefts
A Exposure to chemical contaminants in drinking water in trimesters 1-2
C Volunteering for the case-control study
N_A Exposure to chemical contaminants in drinking water in trimester 3 (impossible by the hypothesized mechanism)

Example 2: Ivers et al. Negative Control Outcome for Selection Bias

Ivers et al.[2] conducted a treatment-facility based case-control study to measure the effectiveness of oral cholera vaccine against cholera in rural Haiti. Investigators were concerned that selective presentation at treatment facilities could bias the measure of effectiveness through selection bias following the structure of Figure 1A. To test whether this bias was present, investigators conducted two parallel case-control studies. The primary study had a case definition of a diarrheal stool that tested positive for cholera (*Y*); the second study used a negative control outcome case definition of a diarrheal stool that tested negative for cholera (N_Y). Both studies used an identical control sampling strategy: they enrolled four geographically matched controls for each case, under the assumption that exposure to cholera vaccine was

geographically clustered. Since the oral cholera vaccine has no protective efficacy against non-cholera diarrhea (an effect that would be impossible by the hypothesized mechanism), any association between cholera vaccine and non-cholera diarrhea would be due to selection bias or other unmeasured confounding. The case-control study estimated the protective efficacy of the vaccine of 58% (13%-80%) against cholera-positive diarrhea and there was no association with non-cholera diarrhea. This lent additional credibility to the study's findings. This general strategy of conducting parallel case-control studies with the auxiliary case-control sample based on a negative control outcome has been used in other vaccine effectiveness studies, including haemophilus influenzae type b (Hib) vaccine.[3]



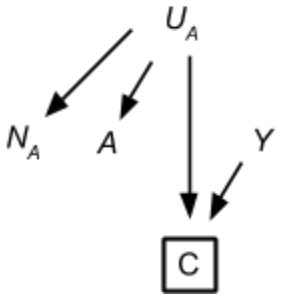
Example 3: de Groot et al. Negative Control Exposure for Selection Bias

A set of case-control studies investigated the association between community-acquired pneumonia and the use of angiotensin-converting-enzyme-inhibitors, statins, and proton pump inhibitors.[4] The investigators hypothesized that selection bias would result if they selected cases from hospitalized pneumonia patients rather than non-hospitalized pneumonia patients. Referral to the hospital is influenced by characteristics like age and comorbidities so patients that are referred and enrolled in the study are likely to be older and have comorbid symptoms; however, these symptoms U_A are also associated with the use of the drugs under study, giving rise to

selection bias ($Y \rightarrow C$ and $U_A \rightarrow C$ through hospital referrals). The investigators included the use of selective serotonin reuptake inhibitors (SSRI) as a negative control exposure N_A that likely shared the same parent factors as the exposures of interest but could not biologically cause pneumonia (Figure 1B). The investigators found significant associations between SSRI use and community-acquired pneumonia, suggesting selection bias could have influenced their other observed associations.

This example illustrates a limitation of negative controls: the assumptions about unobserved variables U encoded in the DAGs are often untestable. In Figure 1B if a directed edge also exists between $U_A \rightarrow Y$, then U_A would also meet the definition of an unmeasured confounder. Bias from unmeasured confounding in addition to selection bias is plausible in the pneumonia example, and illustrates how in practice it may be difficult to distinguish whether bias results from a single structure or from some mixture of them – this limitation surfaces repeatedly in the examples we describe.

Figure 1B (main text) de Groot et al.



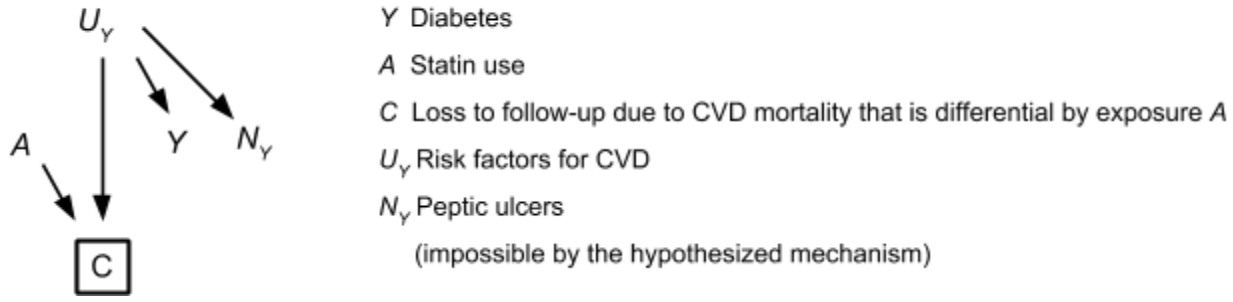
Y Community acquired pneumonia
A Use of ACE inhibitors, statins, & proton pump inhibitors
C Selective enrollment among hospitalized patients
U_A Co-morbidities that cause hospitalization *C* and drug exposures *A*
N_A Selective serotonin reuptake inhibitors
(impossible by the hypothesized mechanism)

Example 4: Danaei et al. Negative Control Outcome for Selection Bias

Danaei et al.[5] used electronic health records to emulate a hypothetical trial on statin use and diabetes, and used a negative control outcome N_Y to detect selection bias from selective loss

to follow-up that could be affected by exposure A (Figure 1C). The investigators hypothesized that because statin use reduces the risk of death from cardiovascular disease, untreated individuals would be more likely to be lost to follow-up due to death from cardiovascular disease (CVD). Furthermore, loss to follow-up would also be affected by an individual's cardiovascular disease risk factors, which are also shared risk factors for diabetes U_Y, thus artificially inflating the risk of diabetes among survivors treated with statins. To detect this bias, investigators included peptic ulcers as a negative control outcome N_Y that is not associated with statins -- an effect that would be impossible by the hypothesized mechanism -- but plausibly shares some of the same risk factors U_Y. The analysis showed no association between statin use and ulcers, lending credibility to the observed association between statins and diabetes.

Figure 1C (main text) Danaei et al.

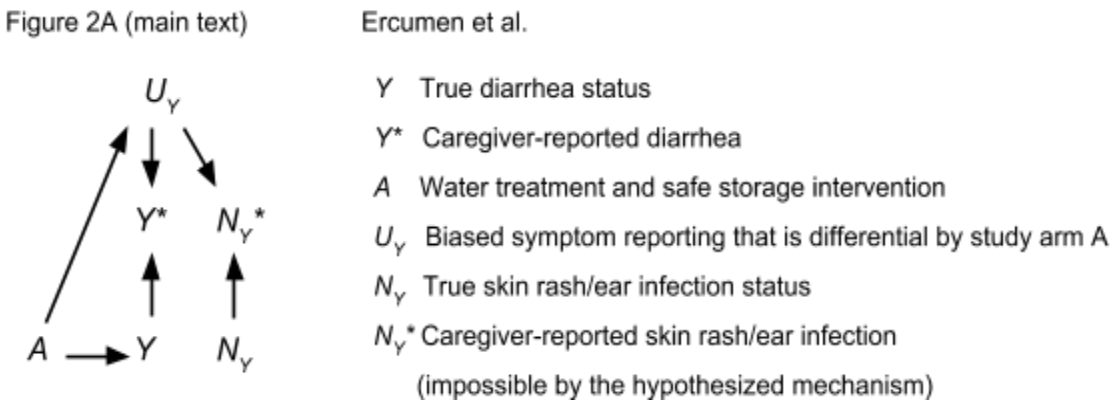


Example 5: Ercumen et al. Negative Control Outcome for Outcome Measurement Error

Ercumen et al.[6] conducted a randomized trial to measure the effect of safe drinking water storage and chlorination on caregiver-reported diarrhea among children used this type of negative control. The intervention reduced diarrhea by 36%, but there was concern that caregiver-reported diarrhea could be subject to courtesy bias if caregivers in the intervention arm differentially under-reported diarrhea to please investigators because the behavior change

component of the intervention emphasized improved child health.[7] If present, the differential measurement error would bias the effect of $A \rightarrow Y^*$ away from the null. To help rule out this possibility, the trial measured caregiver-reported skin rash and ear infections alongside diarrhea symptoms (N_{Y*}), which were thought to be similarly susceptible to differential reporting bias, but could not plausibly be reduced by improved drinking water quality (no direct path from $A \rightarrow N_{Y*}$) -- a check for an effect that would be impossible by the hypothesized mechanism. The trial found no evidence for reductions in either negative control outcome, lending additional credibility to the diarrhea results.

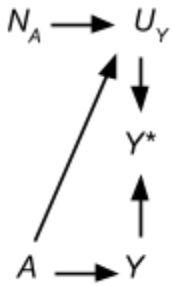
The use of skin rash and ear infections as negative controls in this example assumes that courtesy bias affects negative control reporting N_Y^* in a similar way as it affects diarrhea reporting Y^*; that is, a common source (or correlated sources) of measurement error exists. If instead errors were independent, $U_Y \rightarrow Y^*$ and $\rightarrow N_Y^*$, then it could be possible to see a null result for negative control outcomes and still have bias due to differential reporting of diarrhea. This could occur if participants understood that the water quality intervention was expected to reduce diarrhea, but not ear infections. The example underscores the importance of the assumptions encoded in the DAG -- which are usually unmeasured -- and highlights how negative controls do not necessarily prove the absence of bias when studies "pass" a test (though failing a test is always cause for concern).[8,9]



Example 6: Colford et al. Negative Control Exposure for Outcome Measurement Error

Colford et al.[10] conducted a prospective cohort study to measure the association between *Enterococcus* fecal indicator bacteria in ocean water and incident diarrhea among swimmers at a California beach. The study measured diarrhea using participant-reported symptoms that were potentially subject to reporting errors. Among swimmers who swallowed water, the study estimated that a $\log_{10}$ increase in *Enterococcus* concentration increased the odds of diarrhea by 1.74 (95% CI: 1.25, 2.43). The study also measured incident diarrhea among individuals who were present at the beach, but who did not enter the ocean. By leaving out the essential ingredient -- ingestion of enteric pathogens in ocean water -- there should have been no relationship between *Enterococcus* concentration matched to non-swimmers and incident diarrhea, and indeed that was the case (OR=1.00, 95% CI= 0.76, 1.32). The negative control exposure analysis helped rule out potential sources of unmeasured confounding as well as potential differential reporting errors.

Figure 2A (main text)



Colford et al.

Y True diarrhea status

Y^* Self-reported diarrhea

A Enterococcus concentration in ocean water for swimmers

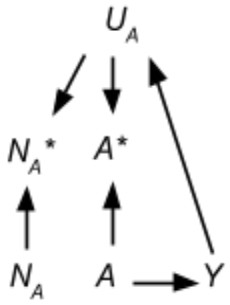
U_Y Biased symptom reporting that is differential by exposure A

N_A Enterococcus concentration in ocean water matched to non-swimmers (leaving out essential ingredient: ingestion of pathogens in ocean water)

Example 7: Zaadstra et al. Negative Control Exposure for Exposure Measurement Error

Zaadstra et al.[11] conducted a retrospective case-control study that measured the association between early childhood illness and multiple sclerosis using this type of negative control exposure. The investigators were concerned that multiple sclerosis cases might remember childhood medical events more accurately than community-based controls, and so in addition to measuring a history of viral infections of interest investigators included in their questionnaire negative control exposures -- broken arm, concussion, and tonsillectomy -- that could not plausibly cause multiple sclerosis. When investigators found significantly elevated odds of multiple sclerosis associated with the negative control exposures, it led to concern about differential recall bias.

Figure 2B (main text)



Zaadstra et al.

Y Multiple sclerosis

A True childhood viral infection status

A^* Reported childhood viral infections

U_A Biased symptom recall that is differential by case status Y

N_A True broken arm, concussion or tonsillectomy in childhood

N_A^* Reported broken arm, concussion or tonsillectomy in childhood (impossible by the hypothesized mechanism)

Example 8. Khush et al. Negative Control Outcome For Exposure Measurement Error

Khush et al.[12] conducted a prospective cohort study in India to measure the association between drinking water quality and child diarrhea. They used a negative control outcome to detect bias from differential exposure measurement error in their analysis (Figure 2B). The analysis matched monthly water samples to a child's diarrhea status at the time of sample collection, with the implicit assumption that water quality measured at the time of diarrhea assessment reflected water quality during the relevant exposure period, 3-10 days earlier. If caregivers boiled water in response to the onset of diarrhea ($Y \rightarrow U_A$), then measured water quality (A^*) would be better than the water quality during the relevant exposure period 3-10 days before diarrhea measurement (A), and this in turn could bias the estimated association toward the null. To help test for this potential bias, the investigators repeated the analysis using cough and congestion/coryza as negative control outcomes, which were unlikely to be caused by fecal water contamination. In this context, caregivers were more likely to boil water if a child had any illness, including cough and congestion/coryza ($N_Y \rightarrow U_A \rightarrow A^*$). The study found an association between water quality and the negative control outcomes. This finding called into question the associations estimated between fecal indicator bacteria concentrations and diarrhea in the study.

Figure 2B (main text)

Khush et al.

- Y Child diarrhea
- A Drinking water quality during relevant period, 3-10 days before diarrhea assessment
- A^* Drinking water quality measured at time of diarrhea assessment
- U_A Water treatment (e.g., boiling) that is differential by outcome Y
- N_Y Cough, congestion/coryza (impossible by the hypothesized mechanism)

ONLINE APPENDIX REFERENCES

1. Ruckart PZ, Bove FJ, Maslia M. Evaluation of exposure to contaminated drinking water and specific birth defects and childhood cancers at Marine Corps Base Camp Lejeune, North Carolina: a case-control study. *Environ Health*. 2013;12:104.

2. Ivers LC, Hilaire IJ, Teng JE, et al. Effectiveness of reactive oral cholera vaccination in rural Haiti: a case-control study and bias-indicator analysis. *Lancet Glob Health*. 2015;3(3):e162-e168.

3. Shapiro ED. Case-control studies of the effectiveness of vaccines: validity and assessment of potential bias. *Pediatr Infect Dis J*. 2004;23(2):127-131.

4. de Groot MCH, Klungel OH, Leufkens HGM, van Dijk L, Grobbee DE, van de Garde EMW. Sources of heterogeneity in case-control studies on associations between statins, ACE-inhibitors, and proton pump inhibitors and risk of pneumonia. *Eur J Epidemiol*. 2014;29(10):767-775.

5. Danaei G, García Rodríguez LA, Fernandez Cantero O, Hernán MA. Statins and risk of diabetes: an analysis of electronic medical records to evaluate possible bias due to differential survival. *Diabetes Care*. 2013;36(5):1236-1240.

6. Ercumen A, Naser AM, Unicomb L, Arnold BF, Colford JM Jr, Luby SP. Effects of source- versus household contamination of tubewell water on child diarrhea in rural Bangladesh: a randomized controlled trial. *PLoS One*. 2015;10(3):e0121907.

7. Schmidt W-P, Cairncross S. Household Water Treatment in Poor Populations: Is There Enough Evidence for Scaling up Now? *Environ Sci Technol*. 2009;43(4):986-992.

8. Lipsitch M, Tchetgen ET, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383-388.

9. Prasad V, Jena AB. Prespecified falsification end points: can they validate true observational associations? *JAMA*. 2013;309(3):241-242.

10. Colford JM, Schiff KC, Griffith JF, et al. Using rapid indicators for Enterococcus to assess the risk of illness after exposure to urban runoff contaminated marine water. *Water Res*. 2012;46(7):2176-2186.

11. Zaadstra BM, Chorus AMJ, van Buuren S, Kalsbeek H, van Noort JM. Selective association of multiple sclerosis with infectious mononucleosis. *Mult Scler*. 2008;14(3):307-313.

12. Khush RS, Arnold BF, Srikanth P, et al. H2S as an Indicator of Water Supply Vulnerability and Health Risk in Low-Resource Settings: A Prospective Cohort Study. *Am J Trop Med Hyg*. 2013;89(2):251-259.