NIH Public Access
Author Manuscript
J Stat Plan Inference. Author manuscript; available in PMC 2010 October 1.

Published in final edited form as:
J Stat Plan Inference. 2009 October 1; 139(10): 3473–3487. doi:10.1016/j.jspi.2009.03.024.

Analytic Bounds on Causal Risk Differences in Directed Acyclic Graphs Involving Three Observed Binary Variables

Sol Kaufman, Ph.D.[1], **Jay S. Kaufman, Ph.D.**[2], and **Richard F. MacLehose, Ph.D.**[3,4]

[1]Department of Otolaryngology, University at Buffalo, 3435 Main Street, Buffalo NY 14214 USA

[2]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, 1020 Pine Ave West, Montreal, Quebec H3A 1A2 Canada

[3]Division of Biostatistics, University of Minnesota, 420 Delaware St SE, Box 303, Minneapolis, MN 55455 USA

[4]Division of Epidemiology, University of Minnesota, 1300 S 2nd St, Suite 300, Minneapolis, MN 55454 USA

Abstract

We apply a linear programming approach which uses the causal risk difference (RD_C) as the objective function and provides minimum and maximum values that RD_C can achieve under any set of linear constraints on the potential response type distribution. We consider two scenarios involving binary exposure X, covariate Z and outcome Y. In the first, Z is not affected by X, and is a potential confounder of the causal effect of X on Y. In the second, Z is affected by X and intermediate in the causal pathway between X and Y. For each scenario we consider various linear constraints corresponding to the presence or absence of arcs in the associated directed acyclic graph (DAG), monotonicity assumptions, and presence or absence of additive-scale interactions. We also estimate Z-stratum-specific bounds when Z is a potential effect measure modifier and bounds for both controlled and natural direct effects when Z is affected by X. In the absence of any additional constraints deriving from background knowledge, the well-known bounds on RDc are duplicated: $-\Pr(Y\neq X) \leq RD_C \leq \Pr(Y=X)$. These bounds have unit width, but can be narrowed by background knowledge-based assumptions. We provide and compare bounds and bound widths for various combinations of assumptions in the two scenarios and apply these bounds to real data from two studies.

Keywords

causality; effect decomposition; confounding; sensitivity analysis; bounding; linear programming; counterfactual models

Address for Correspondence: Jay S. Kaufman, Ph.D., Department of Epidemiology, Biostatistics and Occupational Health, McGill University, 1020 Pine Ave West, Montreal, Quebec H3A 1A2 Canada, Phone: 514-398-7341, Fax: 514-398-4503, e-mail: jay.kaufman@mcgill.ca.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

1. Introduction

This paper develops analytic expressions for deterministic bounds on uncertainties, attributable to unknown confounders, in the observational estimation of causal effects. We demonstrate the importance of background-knowledge-based assumptions in reducing the width of these bounds to useful magnitudes. The scope of our analysis is limited to three measured binary variables and sample size sufficiently large to justify the assumption of negligible sampling error.

The three binary variables, coded 0 for unexposed and 1 for exposed, are designated by: X (exposure or treatment), Y (outcome), and Z (covariate). Their causal relationships are described by two alternative directed acyclic graphs (DAG) in Figures 1 and 2 (Pearl 2000). In Figure 1, Z is a pretreatment variable and potential confounder; we refer to this configuration as the pretreatment-covariate DAG. In Figure 2, Z is an intermediate in the causal chain from X to Y; we refer to this configuration as the intermediate-covariate DAG. The double-headed arrows between all nodes in the graphs denote arbitrary possible confounding of all arcs by unknown confounders. We select the causal risk difference, RD_C, as the measure of average causal effect (ACE) of X on Y.

In the intermediate-covariate DAG case, a particular causal effect measure of interest is the average direct effect of X on Y, i.e., that component of total average causal effect not mediated by Z. Our bounding analysis in this context expands on previous work by Robins and Greenland (1992), Cole and Hernán (2002), and Kaufman, et al. (2005). In the pretreatment-covariate DAG case, Robins (1989) presents some general analytic bounds. Angrist, Imbens, and Rubin (1996) and Balke and Pearl (1997) treat the special case of the non-compliance problem in which Z and X denote *assigned* and *actual* treatments, respectively, the Z→Y arc is missing (i.e., assignment has no direct effect on Y), and Z is randomized (hence, no confounding of the remaining Z→X arc).

The traditional approach to dealing with uncertainty due to confounding-induced non-identifiability of a causal effect is sensitivity analysis (Rothman and Greenland 1998). In sensitivity analyses the possible degree of confounding by unmeasured factors is represented by some set of parameters, for example, those that concern the distribution of a putative confounding variable, its association with exposure/treatment, and its association with the outcome. The analyst applies background knowledge to assign various plausible values, or ranges of plausible values, to these parameters and then quantitatively evaluates the resulting departures of the observational estimate of causal effect from the corrected estimates of causal effect (Flanders and Khoury 1990; Lin, Psaty and Kronmal 1998; Yanagawa 1984).

We adopt an alternative approach in which bounds on RD_C are derived solely on the basis of observational data and the addition of various assumptions concerning plausible response patterns of the population. The framework for this approach is a deterministic counterfactual model in which each individual unit in the population is assumed to have a fixed potential response to each possible input pattern at each node of the DAG having observable inputs, i.e., parent nodes. The observed data can reveal only a single instance of these potential responses. The complete set of potential responses by a particular individual can be thought of as expressing the influences of all factors, both innate and external, acting on that individual. We further assume that the potential responses of each individual are independent of the status of the remaining members of the population. This is the stable-unit-treatment-value assumption (SUTVA) which excludes application to contagious diseases (Rubin 1990).

Each possible pattern of potential responses over all nodes with observed parents is called a “potential response type”. Individuals are categorized by potential response type, and the unknown proportions of these latent potential response types in the population fully define the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

aggregate causal behavior of the population. In this model, RD_C is expressible as a linear function of the potential response type proportions. Moreover, these proportions are subject to linear constraints involving the joint distribution of the observed variables (X, Y, Z). Additional linear constraints on the proportions may also be introduced to reflect background-knowledge-based restrictive assumptions concerning the existence or distribution of plausible causal behaviors within the population.

This system of linear constraints determines a set of *feasible solutions* for the unknown proportions of potential response types, i.e., those proportions which satisfy all the constraints. Within this set of feasible solutions there will be a minimum and a maximum causal risk difference, RD_C; these define the total range of average causal effects consistent with the observed data and imposed assumptions. If the observed data are accurate (i.e., negligible sampling errors and no misclassification or other systemic errors) and the imposed restrictive assumptions are valid, then the true RD_C must lie within this range. The bounds are thus characterized as *deterministic*, as opposed to *probabilistic*. The task of finding the maximum and minimum RD_C within the feasible solution set, under the conditions set forth above, can be accomplished by the computational method of linear programming (LP). This methodology has previously been applied to the pretreatment-covariate DAG, including the specific setting of the non-compliance problem (Balke and Pearl 1997; MacLehose, et al. 2005). It is our intent to expand on these applications and to extend the method to the intermediate-covariate DAG.

In Section 2 we set up the counterfactual model formalism for the pretreatment-covariate DAG and apply a symbolic LP computer program to derive the analytic bounds on causal risk difference under various imposed assumptions. Section 3 does the same for the intermediate-covariate DAG. We provide examples in Section 4 of how these bounds may be applied to real data sets and interpreted. We conclude with a discussion in Section 5 of the characteristics of the bounds and bound widths under various combinations of subject-matter assumptions.

2. Pretreatment-covariate DAG

2. 1 Counterfactual Model Formulation

In the pretreatment-covariate DAG (Figure 1), X (exposure/treatment) and Y (outcome) are nodes with observed parents. The counterfactual or potential response variable for unit u at X is denoted by X_{uz} where index u identifies the individual unit and index z specifies the Z (pretreatment covariate) input (0 or 1)to that unit. Given our deterministic model at the individual unit level, there are exactly four possible patterns of response X_{uz} that unit u can have to input z, viz., $X_{uz}=1$ regardless of z ("doomed"), $X_{uz}=z$ ("causal"), $X_{uz}=1-z$ ("preventive") and $X_{uz}=0$ regardless of z ("immune") (Copas 1973;Greenland and Robins 1986). These four patterns are represented by potential-response-type index values 1, 2, 3, and 4, respectively. Each unit is thus classified as one of these four index values. At node Y the counterfactual or potential response variable for unit u is denoted by Y_{uzx} where, again, u identifies the individual unit, and indices z,x specify the (Z,X) input to that unit. In analogy with the preceding, conditional on the individual unit and conditional on the z input (0 or 1), there are exactly four possible patterns of response Y_{uzx} to input x, viz. (with z fixed) $Y_{uzx}=1$ regardless of x, $Y_{uzx}=x$, $Y_{uzx}=1-x$, and $Y_{uzx}=0$ regardless of x. Thus, each unit, u, is classified as one of $4\times4\times4 = 64$ index triples, [ijk], each index ranging over {1,…,4}, where index i specifies the X_{uz} response pattern, index j specifies the Y_{u0x} response pattern, and index k specifies the Y_{u1x} response pattern. To illustrate, the type [123] refers to a unit in which X would equal 1 regardless of its Z input, Y would equal X if the if the Z value for that unit were 0 (leading to the realized values: Z=0, X=1, Y=1) and Y would equal 1−X if the Z value for that unit were 1 (leading to the realized values: Z=1, X=1, Y=0).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Define q_{ijk} to be the proportion of type [ijk] in the total population. Because Z is generally not independent of the distribution of potential response types (i.e., not ignorable), we need to decompose q_{ijk} into Z=0 and Z=1 components. Define r_{ijk} and s_{ijk} to be the joint proportions of (type = [ijk], Z=0) and (type = [ijk], Z=1), respectively. Then, $q_{ijk} = r_{ijk} + s_{ijk}$. The set $\{r_{ijk}\}$ of all 64 {Z=0} proportions together with the set $\{s_{ijk}\}$ of all 64 {Z=1} proportions completely characterize the causal behavior of the population in the context of the three observed variables (X, Y, Z) and unknown confounding of all arcs in the DAG. The $\{r_{ijk}\}$ and $\{s_{ijk}\}$ cannot be observed directly. What can be known from the data, given the assumption of negligible sampling and classification errors, is the joint distribution of the three observed variables in the total population: p(y,x,z) = Pr(Y=y, X=x, Z=z); y=0,1; x=0,1; z=0,1. The connection between ($\{r_{ijk}\}$, $\{s_{ijk}\}$) and p(y,x,z) is given by eight linear equations, viz., each of the four p(y,x,0) is a sum of 16 of the 64 r_{ijk}, and each of the four p(y,x,1) is a sum of 16 of the 64 s_{ijk} (see Appendix).

The causal risk difference for the total population, RD_C, selected as our objective function to be bounded by application of LP methodology, is defined as: $RD_C = Pr(Y=1|SET[X=1]) - Pr(Y=1|SET[X=0])$, where "SET[X=x]" means that each unit has its X value set to x by external intervention (Pearl 2000, p. 70). Most importantly, RD_C can be expressed as a linear function of the type proportions, specifically, a difference between two partial sums over the $\{r_{ijk}\}$and $\{s_{ijk}\}$ (see Appendix).

2.2 Bounds on ACE Without Additional Background Knowledge

We have just formulated a linear programming (LP) problem, namely, the maximization or minimization of causal risk difference RD_C over the space of feasible values for ($\{r_{ijk}\}$, $\{s_{ijk}\}$), as delimited by the linear constraints involving the observed joint distribution, p(y,x,z), including additional constraints asserting the non-negativity of ($\{r_{ijk}\}$, $\{s_{ijk}\}$) and their overall sum to 1. This problem can be solved numerically by any number of available LP software packages once we are given the observed data for the p(y,x,z). However, our interest is in an analytic solution which expresses the RD_C maximum or minimum as functions of the p(y,x,z) in symbolic form. A program to find such symbolic LP solutions was developed by A. A. Balke in his doctoral dissertation (Balke 1995). This program, OPTIMIZE, accepts a user-defined set of symbolic parameters (in our application, the eight joint probabilities p(0,0,0), …, p(1,1,1)) and requires the objective function and all constraints to be linear in *both* the variables (in our application, r_{ijk} and s_{ijk}) and the symbolic parameters (p(y,x,z)). The program applies LP duality theory (Mattheiss 1973) to find the maximum (minimum) for the objective function expressed as a minimum (maximum) over a set of linear functions of the symbolic parameters. Bound widths are then easily calculated by subtracting the lower bound from the upper bound. OPTIMIZE was made available to us and was run for our problem specification to yield the following lower and upper bounds on RD_C:

$$-[\,p(0,1,0)+p(1,0,0)+p(0,1,1)+p(1,0,1)] \le RD_C \le p(0,0,0)+p(1,1,0)+p(0,0,1)+p(1,1,1)$$

In an equivalent reformulation:

$$-\Pr(Y \neq X) \le RD_C \le \Pr(Y=X)$$

This result is well known (Robins 1989; Manski 1990). The width of these bounds is precisely 1 which, although narrower than [-1,1] (the logical domain of RD_C), is generally not a very

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

useful result and reaffirms the expectation that additional constraints (derived from background knowledge) are required for this method of bounding to have practical utility.

In view of the possibility of effect measure modification by pretreatment covariate Z, it is of interest to find bounds on the stratum-specific causal risk differences, $RD_{C|Z=0}$ and $RD_{C|Z=1}$, defined by:

$$RD_{C\left|Z=z\right.} = \Pr(Y=1 \,|SET[\,X=1\,], Z=z) - \Pr(Y=1|\,SET[\,X=0\,], Z=z);\; z=0, 1$$

Application of the symbolic LP program to this reformulated objective function yields:

$$-\Pr(Y \neq X \left| Z=z\right.) \leq RD_{C\left|Z=z\right.} \leq \Pr(Y=X \left| Z=z\right.);\; z=0, 1$$

In both stratum-specific bounds, the widths remain at 1, although the endpoints may have shifted to reflect possible effect modification.

A further modification in the objective function may be pertinent when covariate Z is experimentally controllable (for example, if Z denotes assignment of treatment X). We define the z-controlled causal risk differences, $RD_{C|SET[Z=0]}$ and $RD_{C|SET[Z=1]}$ by

$$RD_{C\left|SET[\,Z=z]\right.} = \Pr(Y=1 \,|SET[\,X=1\,], SET[\,Z=z] - \Pr(Y=1|\,SET[\,X=0\,], SET[\,Z=z];\; z=0, 1$$

Application of the symbolic LP program to this modified objective function yields:

$$-\Pr(Y \neq X, Z=z) - \Pr(Z \neq z) \leq RD_{C\left|SET[Z=z]\right.} \leq \Pr(Y=X, Z=z) + \Pr(Z \neq z);\; z=0, 1$$

Here, the bounds are wider, the width being given by 1 + Pr(Z≠z). This result is intuitively plausible because in the limiting case of Pr(Z≠z) = 0 no units in the population would be changed by the SET[Z=z] operation; hence, $RD_{C|SET[Z=z]} = RD_{C|Z=z}$. In the opposite limiting case of Pr(Z≠z) = 1, every unit in the population has its natural Z value altered; hence, the observed data provide no useful information, leading to bounds on $RD_{C|SET[Z=z]}$ equal to its full logical domain, [−1, 1].

2.3 Bound Reduction Through Background-Knowledge-Based Assumptions

The following are possible assumptions (expressible as linear constraints) that might be justified on the basis of background knowledge in specific applications. Note that assumptions 1, 2 and 5 were previously considered by Robins (1989) in the particular context of $RD_{C|SET[Z=1]}$.

1. For all units, Z does not affect Y directly (the Z→Y arc in the DAG is absent), i.e., potential response types are restricted to 16 types: [ijj]; i = 1, …, 4; j = 1, …, 4

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2. For all units, Z does not prevent X, and, controlling for Z, X does not prevent Y, i.e., potential response types [3jk], [i3k], and [ij3] do not exist [partial monotonicity].
3. Assumption (2) plus: For all units, controlling for X, Z does not prevent Y, i.e., potential response types restricted to 18 types: [i11], [i22], [i44], [i41], [i42], and [i21]; i = 1, 2, 4 [full monotonicity].
4. Assumption (3) plus: For all units, the joint effects of X and Z on Y are exactly additive (no interaction). Thus, potential response types are restricted to 12 types: [i11], [i22], [i44], and [i41]; i = 1, 2, 4.
5. There are no confounding arcs between Z and X or Y, i.e., Z is exogenous, for example, if Z were physically randomized.

An example where assumption (1) can reasonably be expected to hold is the non-compliance problem, with Z as assigned treatment and X as treatment actually received(Sommer and Zeger 1991). Under assumption (1), RD_C and $RD_{C\,|SET[Z=z]}$ are identical. Assumption (2) asserts monotonicity of Z's effect on X and X's effect on Y. Assumption (3) extends monotonicity to Z's effect on Y. Assumption (4) additionally asserts no interaction between X and Z on Y. Under assumption (5) the potential response type proportions within strata {Z=0} and {Z=1} are equal and identical to those in the total population ($q_{ijk} = r_{ijk} / Pr(Z=0) = s_{ijk} / Pr(Z=1)$). $RD_{C|Z=z}$ and $RD_{C\,|SET[Z=z]}$ are identical in this case also.

The basic set of linear constraints relating $\{r_{ijk}, s_{ijk}\}$ to the observed joint distribution, p(y,x,z), was augmented to represent the above assumptions, separately and in combinations (see Appendix). Where assumption (5) is invoked (Z exogenous), it is convenient to replace the observed data distribution by the joint *conditional* distribution, p(y,x|z) = Pr(Y=y,X=x|Z=z), and it is sufficient to consider only $\{q_{ijk}\}$ in place of $\{r_{ijk}, s_{ijk}\}$. Linear objective functions were formulated for the causal risk differences: RD_C, $RD_{C\,|Z=z}$, and $RD_{C\,|SET[Z=z]}$ (see Appendix). The symbolic LP program results for these various combinations are shown in Table 1.

3. Intermediate-covariate DAG

3. 1 Counterfactual Model Formulation

Development of the counterfactual model for the intermediate-covariate DAG (Figure 2) is similar to that of the pretreatment-covariate DAG (Figure 1) in Section 2: Figure 1 converts to Figure 2 merely by switching the symbols "X" (exposure/treatment) and "Z" (covariate). With the symbols switched, covariate Z (now "intermediate" instead of "pretreatment") and outcome Y become the nodes with observed parents. Each unit, u, is classified as one of 4×4×4 = 64 index triples, [ijk], each representing a potential response type, where index i specifies the Z_{ux} response pattern to x input in unit u, index j specifies the Y_{u0z} response pattern to z input, holding x fixed at 0, and index k specifies the Y_{u1z} response pattern to z input, holding x fixed at 1. Note that type [ijk] holds the same relationship to its DAG structure in both DAGs under consideration. The only difference is in which node is designated as exposure/treatment and which as covariate. This difference comes into play in the definitions of the pertinent RD_C objective functions for the two DAGs.

Continuing, in analogy with the pretreatment-covariate DAG development, we define q'_{ijk} = Pr(response type = [ijk]), r'_{ijk} = Pr(response type = [ijk], X=0), and s'_{ijk} = Pr(response type = [ijk], X=1). Define joint distribution p′(y,z,x) = Pr(Y=y, Z=z, X=x) and joint conditional distribution p′(y,z|x) = Pr(Y=y, Z=z| X=x), noting the change in the order of arguments. The connection between $\{r'_{ijk}, s'_{ijk}\}$ and p′(y,z,x), and similarly that between $\{q'_{ijk}\}$ and p′(y,z|x), is given by a set of eight linear equations *identical in form* to those developed for the pretreatment-covariate DAG (see Appendix).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

We consider the previously defined risk-difference of X on Y in the total population, RD_C, as the average causal effect objective function to be bounded. Because of the explicit presence of intermediate Z, RD_C is distinguished here as the average causal *total* effect. Additionally, there is often interest in bounding the average causal *direct* effect, i.e., that component of average total effect not transmitted through the measured intermediate. The average causal direct effect of X on Y not involving intermediate Z can be defined in two different ways (in risk difference measure):

1. Average *Z-controlled* direct effect (CDE) of X on Y (Kaufman et al., 2005)

$$RD_{C\,|\,SET[Z=z]} = \Pr(Y=1\,|\,SET[X=1], SET[Z=z]) - \Pr(Y=1\,|\,SET[X=0], SET[Z=z]); z=0, 1$$

2. Average *natural* direct effect (NDE) of X on Y (Pearl 2001), also called *pure* direct effect (Robins 2003)

$$RD_{C\,|\,SET(Z=Z_0)} = \Pr(Y=1\,|\,SET[X=1], SET[Z=Z_0]) - \Pr(Y=1\,|\,SET[X=0])$$

where Z_0 is the potential response of Z to SET[X=0]. To clarify this distinction, the CDE is the effect of X on Y within a population whose value of Z is controlled to take on a particular value, whereas the NDE is the effect of X on Y in a population each member of which retains the same value of Z during intervention that it would have had in the absence of intervention (Petersen et al. 2006). As a cautionary note, $RD_{C\,|SET[Z=z]}$ here, although symbolically identical to $RD_{C|SET[Z=z]}$ in Section 2, is conceptually distinct from the latter because of the different roles played by Z in the respective DAGs.

3.2 Bounds on Total and Direct Effects Without Additional Background Knowledge

Our approach is once again structured in the form of an LP problem, namely, the maximization or minimization of one of three linear objective functions: RD_C (total effect), $RD_{C|SET[Z=z]}$ (controlled direct effect), and $RD_{C\,|SET[Z=Z(0)]}$ (natural direct effect) over the space of feasible values for $(\{r'_{ijk}\}, \{s'_{ijk}\})$, as defined by the linear constraints involving the observed joint distribution, $p'(y,z,x)$. These problems were processed by the OPTIMIZE program to yield the following bounds:

$$\begin{array}{lll} -\Pr(Y \neq X) & \leq RD_C & \leq \Pr(Y=X) \\ -\Pr(Y \neq X) - \Pr(Y=X, Z \neq z) & \leq RD_{C\,|\,SET[Z=z]} & \leq \Pr(Y=X)+\Pr(Y \neq X, Z \neq z);\ z=0, 1 \\ -\Pr(Y \neq X) - \Pr(Y=X=1) & \leq RD_{C\,|\,SET[Z=Z_0]} & \leq \Pr(Y=X)+\Pr(Y \neq X=1) \end{array}$$

The corresponding bound widths are: 1, 1 + Pr(Z≠z), and 1 + Pr(X=1), all narrower than the logical domain, [−1, 1], but still too wide to be useful in most applications.

3.3 Bound Reduction Through Background-Knowledge-Based Assumptions

We apply background-knowledge-based assumptions analogous to those considered in Section 2 (with symbols X and Z interchanged to correspond to the shift in applicable DAG):

(1′) For all units, X does not affect Y directly (the X→Y arc in the DAG is absent), i.e., potential response types restricted to 16 types: [ijj]; i = 1, …, 4; j = 1, …, 4. Therefore, there are no controlled or natural direct effects under this assumption.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(2′) For all units, X does not prevent Z, and, controlling for X, Z does not prevent Y, i.e., potential response types [3jk], [i3k], and [ij3] do not exist [partial monotonicity]

(3′) Assumption (2′) plus: For all units, controlling for Z, X does not prevent Y, i.e., potential response types restricted to 18 types: [i11], [i22], [i44], [i41], [i42], and [i21]; i = 1, 2, 4 [full monotonicity].

(4′) Assumption (3′) plus: For all units, the joint effects of X and Z on Y are exactly additive (no interaction). Thus, potential response types are restricted to 12 types: [i11], [i22], [i44], and [i41]; i = 1, 2, 4.

(5′) There are no confounding arcs between X and Z or Y, i.e., X is exogenous, for example, if X were physically randomized.

The basic set of linear constraints relating $\{r'_{ijk}, s'_{ijk}\}$ to the observed joint distribution, $p'(y,x,z)$, was augmented to represent each assumption separately, and in combinations (see Appendix). Linear objective functions were formulated for the total (RD_C), controlled direct ($RD_{C\,|SET[Z=z]}$), and natural direct ($RD_{C\,|SET[Z=Z(0)]}$) causal risk differences (see Appendix). The symbolic LP program results for RD_C, for $RD_{C\,|SET[Z=z]}$ and for $RD_{C\,|SET[Z=Z(0)]}$ are shown in Table 2.

4. Examples

While we have pursued simulated data examples in previous papers (Kaufman et al, 2005), we do not do so here. The bounds produced by this method are deterministic so there is little be demonstrated in simulated data. Instead, we focus on two data analytic examples to demonstrate both inference and interpretation.

4.1 Pretreatment-covariate DAG

Brookhart et al. conducted an instrumental variable (IV) analysis to estimate the causal effect of treatment with COX-2 inhibitors relative to nonselective, non-steroidal anti-inflammatory medications (NSAIDs) on occurrence of gastrointestinal complications within 60 days (Brookhart et al. 2006). The IV was the prescribing physician's time-varying "preference" for treatment with COX-2 inhibitors or nonselective NSAIDs, as estimated from their previous prescription. With physicians' preference variable labeled as Z, exposure labeled as X and outcome labeled as Y, the DAG for this problem conforms to the pretreatment-covariate DAG shown in Figure 1. Using this approach, the authors found a protective effect of COX-2 exposure (risk difference = 0.92 / 100 patients).

We coded Z=1 to represent a physicians preference for nonselective NSAIDs, X = 1 to represent a prescription for nonselective NSAIDs, and Y = 1 to represent an outcome of gastrointestinal complications. In a total of 49919 prescriptions studied, approximately 65% were for nonselective NSAIDs, and the outcome occurred in only 321 (0.6%) subjects. The crude effect of exposure on outcome was estimated as $Pr(Y=1|X=1) - Pr(Y=1|X=0) = -0.03$ / 100 patients, whereas the crude effect of the instrument on the outcome was $Pr(Y=1|Z=1) - Pr(Y=1|Z=0) = 0.210$ / 100 patients.

Making no additional causal assumptions, the bounds on the causal risk difference (RD_C) of the effect of exposure on outcome in the total population are $-Pr(Y \neq X) \leq RD_C \leq Pr(Y = X)$, or [−0.30, 0.70] and with unit length are therefore relatively uninformative. Adding only assumption (1), the "exclusion restriction" of no direct effect of Z on Y, does not narrow the bounds further (Table 1). Bounds in the Z=0 sub-group are [−0.23, 0.77], and bounds in the Z=1 sub-group are [−0.46, 0.54]. It seems plausible to assume that physician preference for one or the other drug might have an effect only through the writing of a prescription, and this assumption is crucial for the IV method to be valid. But it is also possible that unmeasured

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

physician preferences might influence not only choice of drug but also other aspects of care that affect outcome (Hernan and Robins 2006).

We have coded the variables to be consistent with monotonicity of effects, justifying assumptions (2) and (3) of "partial monotonicity" and "full monotonicity", respectively. Applying these assumptions narrows the bounds mostly by truncating at 0, so that they become [0, 0.77] in the Z = 0 stratum and [0, 0.54] in the Z = 1 stratum. Additional imposition of assumption 4 "no Z-X interaction", which is also invoked in standard IV analyses, reduces the bounds to the minimum under the previous assumptions, or [0, 0.54].

Assuming only exogeneity of Z (assumption (5)) does not narrow the bounds from those under no assumptions. However, adding both assumptions (1) and (5) jointly, the absence of a direct effect of Z on Y and the exogeneity of Z, yields the Balke-Pearl bounds of [−0.23, 0.54], having width of 0.77. If the usual assumption of exogeneity of Z is met but there is a direct effect of Z on the outcome, then the standard IV method is biased, but we may still rely on bounds for other combinations of assumptions. For example, if Z is unconfounded and the effects are monotonic, then one could invoke assumptions 2 and 5 jointly, or 3 and 5 jointly in the case of additional monotonicity of the Z effect on Y. This leads to bounds of [0, 0.77] and [0, 0.54] in the Z = 0 and Z = 1 subgroups, respectively, under assumptions (2) and (5), and to bounds of [0, 0.54] for both subgroups under assumptions (3) and (5). This latter set of bounds remains the same under assumptions (4) and (5). Alternatively, if assumption (1) of no direct effect of the instrument is valid, and if monotonicity can additionally be assumed, but Z may be confounded (i.e. assumption (5) violated), then the bounds remain as [0, 0.77] and [0, 0.54] in the Z = 0 and Z = 1 subgroups, respectively, or [0, 0.54] for both subgroups under assumptions (1) and (4).

If all the usual structural assumptions for the IV analysis are met, the causal effect of exposure X is still not identified as a point estimate (Hernan and Robins 2006). To obtain a point estimate, one must restrict to local (i.e., sub-population-specific) inference. For the entire study population, however, bounds under all assumptions jointly (i.e., assumptions (1,2,5) or (1,3,5) or (1,4,5)) are [0.002, 0.54] for both Z = 0 and Z = 1 groups, with width of 0.54. The authors' IV point estimate for the effect in the complier sub-population was 0.0092, close to the lower bound of this interval.

In this particular application, the addition of assumptions narrows the bounds from unit length to roughly half that width, but the resulting interval may still be too wide to exclude values of potential interest. That is, the RD_C bounds under all the structural assumptions considered here are [0, 0.54], but subject matter considerations may already be sufficient to exclude RD_C values above 0.5 anyway. This reveals that in the present example, the observed data table and the additional assumptions invoked are consistent with a range of average causal effects that includes essentially all biologically plausible values. This could be seen as a potential drawback of the non-parametric analysis approach described in this paper, since relatively weak effects, which are common in real-world applications, will often generate relatively wide bounds, and therefore fail to contribute much to inference beyond what can be gleaned from subject matter considerations alone. On the other hand, this can be seen as a strength of the method, since it reveals how much uncertainty is truly present in the data, and is reduced to a point estimate only through the application of strong assumptions that are not verifiable in the data themselves.

4.2 Intermediate-covariate DAG

In 2003, Costello and colleagues reported on the analysis of a natural experiment involving income supplementation (Costello et al. 2003). Their data came from an 8-year longitudinal study of the development of psychiatric disorders in 1420 rural and urban youth from 11 counties in western North Carolina, with an over-sample of Native American (Cherokee)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

children. Beginning half-way through the follow-up period, tribal members began to receive income from a gambling casino that opened on a reservation in the study area.

The authors calculated mean 4-year psychiatric symptom scores for the period before the casino opened (1993-1996) and for the 4-year period after the casino opened (1997-2000). Families were defined as poor if the mean income for the 4-year period, adjusted for family size, was below the federal poverty line. Families then were classified into 3 groups: (1) persistently poor, those families below the federal poverty line before and after the casino opened; (2) ex-poor, those families who moved out of poverty after the casino opened; and (3) never poor, those families above the poverty line before and after the casino opened. The authors sought to estimate the causal relation between alleviation of poverty (X) and child psychopathology (Y) when exposure could be considered exogenous. In addition to the estimation of a total effect of poverty alleviation, the authors also sought to estimate the proportion of the effect that was attributable to a measured intermediate, namely “lax supervision” of the child by the parents (Z). Controlling for the intermediate in a multiple regression model, the authors reported that 77% of the total effect was attributable to this intermediate.

After the Casino opening, 14% of Cherokee families in the study rose out of poverty, while 53% remained poor and 32% were never poor. Non-Indian families were unaffected by the casino. The measured intermediate variable, inadequate parental supervision, was defined as the inability to exercise control over a child at least once per week. The prevalence of inadequate parental supervision was roughly 19%, changing from 23% before the casino to 12% after the casino. The outcome was originally measured as a count variable for the number of psychiatric symptoms, but here we analyze the outcome as a dichotomization between 0-2 psychiatric symptoms versus 3 or more psychiatric symptoms. The prevalence of the outcome defined by this categorization was roughly 37%, changing from 47% before the casino opened to 16% after the casino opened.

This scenario therefore corresponds to the intermediate-covariate DAG shown in Figure 2. Coding all deleterious states as exposed (i.e., $X = 1$ represents poverty, $Z = 1$ represents inadequate parental control and $Y = 1$ represents a higher count of psychiatric symptoms), the estimated average causal effect of exposure X on outcome Y is then $\Pr(Y=1|X=1) - \Pr(Y=1|X=0) = 0.31$. The estimated average causal effect of intermediate Z on outcome Y is $\Pr(Y=1|Z=1) - \Pr(Y=1|Z=0) = 0.53$, and the estimated average causal effect of exposure X on intermediate Z is $\Pr(Z=1|X=1) - \mathrm{PR}(Z=1|X=0) = 0.11$. Analytic solutions for bounds for various causal effects of interest are shown in Table 2. Bounds on the average causal effect of exposure X on outcome Y are defined by $-\Pr(Y \neq X) \leq \mathrm{RD}_C \leq \Pr(Y = X)$, or $[-0.41, 0.59]$ and with unit length are therefore relatively uninformative. Addition of assumption (2′) (partial monotonicity) is quite plausible in this example, but alone does not narrow the bounds. Addition of assumption (3′) (full monotonicity) is also quite plausible, and narrows the bounds by truncating the lower bound, $0 \leq \mathrm{RD}_C \leq \Pr(Y = X)$, or $[0, 0.59]$.

The uniquely attractive and fortuitous quality of the natural experiment, however, is that the exposure is plausibly exogenous in this case, justifying assumption (5′). This is because the period before and after the advent of the casino can be considered exchangeable with respect to all other aspects of the subjects and their environments. This therefore narrows the bounds of the total average causal effect of X on Y to the observed point estimate = 0.31. Uncertainty still remains in the effect decomposition with respect to measured intermediate Z, however, since the quasi-randomization of X does not in any way suggest an absence of confounding in the relation between Z and Y, a point that was not acknowledged by the authors of the original publication. Attention therefore shifts to the controlled and natural direct effects of poverty alleviation exposure (X) on outcome, with respect to the measured intermediate representing inadequate parental control (Z).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Without any additional assumptions, bounds on the direct effects are also uselessly wide. The stratum-specific controlled direct effect bounds for $RD_{C|SET[Z=z]}$ are [−0.55, 0.64] for Z = 0 and [−0.86, 0.95] for Z = 1. Likewise, bounds for the natural direct are [−0.73, 0.95], considering the non-poor (X=0) as the reference exposure level. Adding the plausible assumption of full monotonicity (assumption (3′)) reduces bound width considerably and yields controlled direct effect bounds of [0, 0.62] and [0, 0.92] for the Z = 0 and Z = 1 strata, respectively. Under the full monotonicity assumption, the bounds on the natural direct effect are [0, 0.59], probably too wide to be very informative.

Once again, however, the unique advantage of the natural experiment is that the exogeneity assumption can be asserted, which narrows the direct bounds. The controlled direct effect bounds under assumption (5′) are [0.08, 0.43] for the Z = 0 stratum and [−0.77, 0.88] for the Z = 1 stratum. Likewise, the natural direct effect bounds under assumption (5′) are [0.003, 0.47]. We can do better than this, however, because the (full) monotonicity assumption (3′) is also quite reasonable in this example, leading to controlled direct effect bounds of [0.20, 0.39] and [0.002, 0.80] in the Z = 0 and Z = 1 strata, respectively. The natural direct effect bounds under assumptions (3′,5′) are [0.21, 0.31]. An example interpretation of these bounds is that the estimated controlled direct effect $RD_{C|SET[Z=0]}$ is the causal effect of moving out of poverty on psychiatric symptoms if all families were constrained to have adequate parental supervision (e.g., by provision of universally subsidized childcare). The estimated value of this effect is 0.27. Under assumptions of monotonicity and exogeneity which are justified on subject matter grounds, the deterministic bounds on $RD_{C|SET[Z=0]}$ are [0.20, 0.39]. Assuming no sampling variability and assuming no other improprieties in the study (e.g., misclassification of exposure or outcome), the causal effect must fall within these bounds, based on the observed table and additional assumptions (3′,5′). For the natural direct effect, the interpretation would be the causal effect of moving out of poverty on psychiatric symptoms if all families were constrained to have level of parental supervision that they would have had if they were not poor.

The authors reported a single uniform estimate for the (controlled) direct effect, since they used a regression model, and so in order to make our estimate more directly comparable to that reported in the original study, we also invoke homogeneity assumption (4′), which prohibits additive scale interaction between X and Z. This forces the Z-stratum-specific average causal direct effects to be equal. Under these assumptions, i.e., (3′,4′,5′), the bounds on both the average controlled direct effect and the average natural direct effect are [0.21, 0.31]. The point estimate for this single uniform effect, standardized to the total study population, is 0.25. This is the estimated causal effect of moving out of poverty on psychiatric symptoms if all families were constrained to have the same level of parental supervision. The authors' published assertion is that (100-77) = 33% of the total effect is not accounted for by the measured intermediate Z. Our estimate, expressed as a proportion of the estimated total average causal effect, is 0.25/0.31 = 81%. Inverting our bounds and expressing them as a proportion of the total average causal effect of X on Y, we obtain bounds for this quantity of interest of [0.21/0.31 = 68%, 0.31/0.31 = 100%]. Albeit using slightly different effect measure and outcome definitions, these results are not consistent with the published findings. Given validity of assumptions and no other improprieties, the average causal direct effect must lie between 68% and 100% of the total effect, excluding the value of 33% reported by the authors.

5. Discussion

5.1 Pretreatment-covariate DAG (Table 1)

Assumption (1) "Z does not affect Y directly" and assumption (5) "Z exogenous", each separately narrow the bounds on $RD_{C|SET[Z=z]}$ somewhat. The main effect of the monotonicity assumptions, (2) and (3), is to force a lower bound of zero on all three causal effect objective functions. Additionally, the upper bound on $RD_{C|SET[Z=z]}$ is progressively reduced as

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

assumption (3) “full mononoticity”, and then assumption (4) “no Z-X interaction” are imposed. Combinations of these assumptions lead to further narrowing of bounds that could be of practical utility. In particular, (1)+(5) yield the Balke-Pearl bounds, which were originally derived using the same OPTIMIZE program (Balke and Pearl 1997), and (2)+(5), (3)+(5) and (4)+(5) yield progressively further reductions in the upper bound for some or all of the three causal effect objective functions. Under assumption (1), the monotonicity/no-interaction assumptions (2), (3) and (4) are indistinguishable, since the added restrictions in (3) and (4) are rendered vacuous by the absence of a direct Z→Y arc. If assumption (5) “Z exogenous” is also imposed, further narrowing is achieved and the lower bound can even be strictly positive.

Our results in Table 1 for z-controlled causal risk differences, $RD_{C|SET[Z=z]}$, may be compared with those in Table 1 of Robins (1989). The relation between Robins’ and our assumptions is as follows: Robins’ (1) is equivalent to our (1) “Z does not affect Y directly”; Robins’ (2)+(4) are equivalent to our (2) “partial monotonicity”; Robins’ (3) is equivalent to our (5) “Z exogenous”. We have no counterparts to Robins’ assumptions (2) or (4), individually, nor to his assumptions (5) through (8), individually or in any combination. Robins has no counterpart to our assumptions (4) or (5). Because of the limited correspondence in assumptions, comparisons can only be made with Rows 1, 2, 5, and 6 in Robins’ Table 1. Algebraic reformulation of Robins’ bounds (with correction of two obvious typographical errors) reveals agreement on bounds given in Row 1 (no assumptions), Row 2 (Z does not affect Y directly), and Row 6 (Z exogenous), but not for the bounds given in Row 5. Row 5 corresponds to the case analyzed by Balke and Pearl, for which our results are in agreement. Robins’ results in Row 5 differ, apparently because of his assumption of weak randomization of treatment assignment, in place of our and Balke and Pearl’s assumption of strong randomization (Robins and Greenland 1996). The remaining bounds in Table 1 have not to our knowledge been published previously.

In connection with our results, we note an important property. Assumption (5), alone, and, alternatively, any combination of the other four assumptions are logically compatible with *any* observed joint distribution, p(y, x, z), because of the presence of unknown confounding allowed by these assumptions. On the other hand, (5) plus one or more of the other assumptions imply the existence of incompatible joint probability distributions which, if observed, could serve to disconfirm the asserted assumptions.

The results in Table 1, having been derived through a symbolic linear programming procedure, can be applied to actual observational data to compute bounds without the need for the analyst to actually do an LP optimization; one merely substitutes observed values for p(y,x,z). However, having LP software available could be useful in the case that background-knowledge assumptions restrict the domain of p(y, x, z). In such instances, running the observed joint distribution together with the assumptions through an LP program would test whether the distribution is in the admissible domain. If it is not, the program would halt with a “no feasible solution” output, and that would raise a flag concerning the validity of the imposed assumptions.

5.2 Intermediate-covariate DAG (Table 2)

The bounds in Table 2 have not to our knowledge been published previously. In particular, formalization of natural direct effects did not occur until 2001 (Pearl 2001), and therefore these quantities were not investigated earlier (although natural direct effects are described qualitatively in Robins & Greenland 1992).

It is seen in Table 2 that total causal effect bounds are not narrowed by either assumption (1′) “X does not affect Y directly” or (2′) “partial monotonicity”, alone, a result that is intuitive. On the other hand, “full monotonicity”, with or without “no X-Z interaction”, (3′) or (4′), brings

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the lower bound up to zero, and, interestingly, the combination of (1′) with any of (2′), (3′) or (4′) can lead to further narrowing which comes from inclusion of the measured intermediate covariate data. As expected, invocation of (5′), "X exogenous", with or without any other assumption(s) renders the total causal effect of X on Y determined by the observed data; hence bound width reduces to zero.

We next consider the two direct causal effect measures, z-controlled direct effect and natural direct effect. Here, assumption (1′) "X does not affect Y directly" is obviously not relevant. We see that assumption (2′) "partial monotonicity", while it can produce some narrowing of direct causal effect bounds, does not lead to a zero lower bound. That is achieved with assumption (3′) "full monotonicity" which, when combined with (4′) "no X-Z interaction" leads to progressively narrower bounds. But, as in the pretreatment configuration, the best one can do with these monotonicity/no-interaction assumptions alone is the bound [0, Pr(Y=X)]. Controlled direct effect bounds are wider than natural direct effect bounds under assumption (3′), because of the Pr(Y≠X, Z≠X) term in the former, but this difference disappears when assumption (4′) is also invoked. Assumption (5′) "X exogenous", with or without additional assumptions, yield more complicated bounds on both direct causal effect measures, and the circumstances under which there could be substantial narrowing of bound width need to be assessed through specific numerical examples. One possible consequence of exogenous X together with monotonicity of effects is the existence of strictly, and even substantively, positive lower bounds.

As in the case of the pretreatment-covariate DAG, assumption (5′) plus at least one of the other assumptions limits the domain of joint distributions of the observed variables, so that realization of a prohibited distribution could serve to disconfirm the selected set of assumptions.

Appendix: Linear Program Formulation Details

A.1 Pretreatment-Covariate DAG – No Added Assumptions

Z →X→Y, Z →Y, as in Figure 1. All three nodes are binary, coded 0 or 1.

Potential Response Types: [ijk]; i=1,…,4; j=1,…,4; k=1,…,4

Index i defines potential response X_{uz} of unit u to input z at node X

Index j defines potential response Y_{u0x} of unit u to input x at node Y when input z = 0

Index k defines potential response Y_{u1x} of unit u to input x at node Y when input z = 1

index = 1 denotes potential response = 1, regardless of input

index = 2 denotes potential response = input

index = 3 denotes potential response = 1 − input

index = 4 denotes potential response = 0, regardless of input

Variables: r_{ijk}, s_{ijk} (alternatively, q_{ijk}); i=1,…,4; j=1,…,4; k=1,…,4

r_{ijk} = Pr(potential response type = [ijk], Z=0)

s_{ijk} = Pr(potential response type = [ijk], Z=1)

q_{ijk} = Pr(potential response type = [ijk]) [Note: $q_{ijk} = r_{ijk} + s_{ijk}$]

Dot notation will be used for summation, e.g., $r_{i\cdot k} = r_{i1k} + r_{i2k} + r_{i3k} + r_{i4k}$

Parameters (observed data):

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$p(y, x, z) = Pr(Y=y, X=x, Z=z)$; y=0,1; x=0,1; z=0,1

Alternatively, $p(y, x| z) = Pr(Y=y, X=x |Z=z)$; y=0,1; x=0,1; z=0,1

Basic Linear constraints (without further assumptions):

$p(0,0,0) = r_{22\cdot} + r_{24\cdot} + r_{42\cdot} + r_{44\cdot}$

$p(0,1,0) = r_{13\cdot} + r_{14\cdot} + r_{33\cdot} + r_{34\cdot}$

$p(1,0,0) = r_{21\cdot} + r_{23\cdot} + r_{41\cdot} + r_{43\cdot}$

$p(1,1,0) = r_{11\cdot} + r_{12\cdot} + r_{31\cdot} + r_{32\cdot}$

$p(0,0,1) = s_{3\cdot2} + s_{3\cdot4} + s_{4\cdot2} + s_{4\cdot4}$

$p(0,1,1) = s_{1\cdot3} + s_{1\cdot4} + s_{2\cdot3} + s_{2\cdot4}$

$p(1,0,1) = s_{3\cdot1} + s_{3\cdot3} + s_{4\cdot1} + s_{4\cdot3}$

$p(1,1,1) = s_{1\cdot1} + s_{1\cdot2} + s_{2\cdot1} + s_{2\cdot2}$

$r\ldots + s\ldots = 1$

$r_{ijk} \geq 0$

$s_{ijk} \geq 0$

Objective functions:

Total Causal Effect of X on Y:

$RD_C = r_{\cdot2\cdot} + s_{\cdot\cdot2} - r_{\cdot3\cdot} - s_{\cdot\cdot3}$

Z-Stratified Causal Effects of X on Y:

$RD_{C\,|Z=0} = (r_{\cdot2\cdot} - r_{\cdot3\cdot}) / Pr(Z=0)$

$RD_{C\,|Z=1} = (s_{\cdot\cdot2} - s_{\cdot\cdot3}) / Pr(Z=1)$

Z-Controlled Causal Effects of X on Y:

$RD_{C\,|SET[Z=0]} = r_{\cdot2\cdot} + s_{\cdot2\cdot} - r_{\cdot3\cdot} - s_{\cdot3\cdot}$

$RD_{C\,|SET[Z=1]} = r_{\cdot\cdot2} + s_{\cdot\cdot2} - r_{\cdot\cdot3} - s_{\cdot\cdot3}$

A.1.1 Augmentation of Constraints to Reflect Background-Knowledge-Based Assumptions

Assumption (1) [Z does not affect Y directly]

$r_{ijk} = s_{ijk} = 0$ if $j \neq k$

Assumption (2) [Z does not prevent X and X does not prevent Y, controlling for Z, i.e., partial monotonicity]

$r_{3jk} = s_{3jk} = r_{i3k} = s_{i3k} = r_{ij3} = s_{ij3} = 0$

Assumption (3) [(2) plus Z does not prevent Y, controlling for X, i.e., full monotonicity]

All r_{ijk} and s_{ijk} variables constrained to zero with the exception of r_{i11}, r_{i22}, r_{i44}, r_{i41}, r_{i42}, r_{i21}, s_{i11}, s_{i22}, s_{i44}, s_{i41}, s_{i42}, s_{i21}; i=1, 2, 4

Assumption (4) [(3) plus X and Z do not interact at Y]

All r_{ijk} and s_{ijk} variables constrained to zero with the exception of r_{i11}, r_{i22}, r_{i44}, r_{i41}, s_{i11}, s_{i22}, s_{i44}, s_{i41} ; i=1, 2, 4

Assumption (5) [Z exogenous]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Potential response type proportions are the same in both Z strata, {Z=0}and {Z=1}. This implies that $r_{ijk}/\Pr(Z=0) = s_{ijk}/\Pr(Z=1) = q_{ijk}$. Use q_{ijk} in place of r_{ijk} and s_{ijk} as the LP variables to be solved for objective function optimization. Replace the parameters p(y, x, z) by p(y, x| z).

The eight equality constraints then become

$$p(0,0|\,0) = q_{22\cdot} + q_{24\cdot} + q_{42\cdot} + q_{44\cdot}$$

$$p(0,1|\,0) = q_{13\cdot} + q_{14\cdot} + q_{33\cdot} + q_{34\cdot}$$

$$p(1,0|\,0) = q_{21\cdot} + q_{23\cdot} + q_{41\cdot} + q_{43\cdot}$$

$$p(1,1|\,0) = q_{11\cdot} + q_{12\cdot} + q_{31\cdot} + q_{32\cdot}$$

$$p(0,0|\,1) = q_{3\cdot2} + q_{3\cdot4} + q_{4\cdot2} + q_{4\cdot4}$$

$$p(0,1|\,1) = q_{1\cdot3} + q_{1\cdot4} + q_{2\cdot3} + q_{2\cdot4}$$

$$p(1,0|\,1) = q_{3\cdot1} + q_{3\cdot3} + q_{4\cdot1} + q_{4\cdot3}$$

$$p(1,1|\,1) = q_{1\cdot1} + q_{1\cdot2} + q_{2\cdot1} + q_{2\cdot2}$$

The objective functions are modified to:

$$RD_C = (q_{\cdot2\cdot} - q_{\cdot3\cdot})\Pr(Z=0) + (q_{\cdot\cdot2} - q_{\cdot\cdot3})\Pr(Z=1)$$

$$RD_{C\,|Z=0} = RD_{C\,|SET[Z=0]} = q_{\cdot2\cdot} - q_{\cdot3\cdot}$$

$$RD_{C\,|Z=1} = RD_{C\,|SET[Z=1]} = q_{\cdot\cdot2} - q_{\cdot\cdot3}$$

A.2 Intermediate-Covariate DAG – No Added Assumptions

X →Z →Y, X →Y, as in Figure 2. All three nodes are binary, coded 0 or 1. [Identical to pretreatment-covariate DAG with X and Z interchanged]

Potential Response Types: [ijk]; i=1,…,4; j=1,…,4; k=1,…,4

Index i defines potential response Z_{ux} of unit u to input x at node Y

Index j defines potential response Y_{u0z} of unit u to input z at node Y when input x = 0

Index k defines potential response Y_{u1z} of unit u to input z at node Y when input x = 1

index = 1 denotes potential response = 1, regardless of input

index = 2 denotes potential response = input

index = 3 denotes potential response = 1 − input

index = 4 denotes potential response = 0, regardless of input

Variables: r'_{ijk}, s'_{ijk} (alternatively, q'_{ijk}); i=1,…,4; j=1,…,4; k=1,…,4

r'_{ijk} = Pr(potential response type = [ijk], X=0)

s'_{ijk} = Pr(potential response type = [ijk], X=1)

q'_{ijk} = Pr(potential response type = [ijk]) [Note: $q'_{ijk} = r'_{ijk} + s'_{ijk}$]

Dot notation will be used for summation, e.g., $r'_{i\cdot k} = r'_{i1k} + r'_{i2k} + r'_{i3k} + r'_{i4k}$

Parameters (observed data):

$p'(y, z, x) = \Pr(Y=y, Z=z, X=x)$; y=0,1; z=0,1; x=0,1

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Alternatively, $p'(y, z| x) = Pr(Y=y, Z=z |X=x); y=0,1; z=0,1; x=0,1$

Basic Linear constraints (without further assumptios):

Because the pretreatment-covariate DAG and intermediate-covariate DAG are identical except for symbolic interchange of Z and X, the linear constraints involving $p'(y, z, x)$, r'_{ijk}, and s'_{ijk} are identical in form to those in Section A.1 involving $p(y, x, z)$, r_{ijk}, and s_{ijk}., respectively.

Objective functions:

Total Causal Effect of X on Y:

$$\begin{aligned} RD_C = \; & r'_{131}+r'_{132}+r'_{141}+r'_{142}+r'_{221}+r'_{222}+r'_{241}+r'_{242}+ \\ & r'_{331}+r'_{333}+r'_{341}+r'_{343}+r'_{421}+r'_{423}+r'_{441}+r'_{443}- \\ & r'_{113}-r'_{114}-r'_{123}-r'_{124}-r'_{213}-r'_{214}-r'_{233}-r'_{234}- \\ & r'_{312}-r'_{314}-r'_{322}-r'_{324}-r'_{412}-r'_{414}-r'_{432}-r'_{434}+ \\ & s'_{131}+s'_{132}+s'_{141}+s'_{142}+s'_{221}+s'_{222}+s'_{241}+s'_{242}+ \\ & s'_{331}+s'_{333}+s'_{341}+s'_{343}+s'_{421}+s'_{423}+s'_{441}+s'_{443}- \\ & s'_{113}-s'_{114}-s'_{123}-s'_{124}-s'_{213}-s'_{214}-s'_{233}-s'_{234}- \\ & s'_{312}-s'_{314}-s'_{322}-s'_{324}-s'_{412}-s'_{414}-s'_{432}-s'_{434} \end{aligned}$$

Controlled Direct Effect of X on Y:

$$\begin{aligned} RD_{C\,|SET[Z=0]} = \; & r'_{\cdot 21}+r'_{\cdot 23}+r'_{\cdot 41}+r'_{\cdot 43}-r'_{\cdot 12}-r'_{\cdot 14}-r'_{\cdot 32}-r'_{\cdot 34}+ \\ & s'_{\cdot 21}+s'_{\cdot 23}+s'_{\cdot 41}+s'_{\cdot 43}-s'_{\cdot 12}-s'_{\cdot 14}-s'_{\cdot 32}-s'_{\cdot 34} \\ RD_{C\,|SET[Z=1]} = \; & r'_{\cdot 31}+r'_{\cdot 32}+r'_{\cdot 41}+r'_{\cdot 42}-r'_{\cdot 13}-r'_{\cdot 14}-r'_{\cdot 23}-r'_{\cdot 24}+ \\ & s'_{\cdot 31}+s'_{\cdot 32}+s'_{\cdot 41}+s'_{\cdot 42}-s'_{\cdot 13}-s'_{\cdot 14}-s'_{\cdot 23}-s'_{\cdot 24} \end{aligned}$$

Natural Direct Effect of X on Y:

$$\begin{aligned} RD_{C\,|SET[Z=Z(0)]} = \; & r'_{131}+r'_{132}+r'_{141}+r'_{142}+r'_{221}+r'_{223}+r'_{241}+r'_{243}+ \\ & r'_{331}+r'_{332}+r'_{341}+r'_{342}+r'_{421}+r'_{423}+r'_{441}+r'_{443}- \\ & r'_{113}-r'_{114}-r'_{123}-r'_{124}-r'_{212}-r'_{214}-r'_{232}-r'_{234}- \\ & r'_{312}-r'_{314}-r'_{323}-r'_{324}-r'_{412}-r'_{414}-r'_{432}-r'_{434}+ \\ & s'_{131}+s'_{132}+s'_{141}+s'_{142}+s'_{221}+s'_{223}+s'_{241}+s'_{243}+ \\ & s'_{331}+s'_{332}+s'_{341}+s'_{342}+s'_{421}+s'_{423}+s'_{441}+s'_{443}- \\ & s'_{113}-s'_{114}-s'_{123}-s'_{124}-s'_{212}-s'_{214}-s'_{232}-s'_{234}- \\ & s'_{312}-s'_{314}-s'_{323}-s'_{324}-s'_{412}-s'_{414}-s'_{432}-s'_{434} \end{aligned}$$

A.2.1 Augmentation of Constraints to Reflect Background-Knowledge-Based Assumptions

Assumption (1′) [X does not affect Y directly]

$q'_{ijk} = r'_{ijk} = s'_{ijk} = 0$ if $j \neq k$

Assumption (2′) [X does not prevent Z and Z does not prevent Y, controlling for X, i.e., partial monotonicity]

$r'_{3jk} = s'_{3jk} = r'_{i3k} = s'_{i3k} = r'_{ij3} = s'_{ij3} = 0$

Assumption (3′) [(2′) plus X does not prevent Y, controlling for Z, i.e., total monotonicity]

All r'_{ijk} and s'_{ijk} variables constrained to zero with the exception of r'_{i11}, r'_{i22}, r'_{i44}, r'_{i41}, r'_{i42}, r'_{i21}, s'_{i11}, s'_{i22}, s'_{i44}, s'_{i41}, s'_{i42}, s'_{i21}; i=1, 2, 4

Assumption (4′) [(3′) plus X and Z do not interact at Y]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

All r'_{ijk} and s'_{ijk} variables constrained to zero with the exception of r'_{i11}, r'_{i22}, r'_{i44}, r'_{i41}, s'_{i11}, s'_{i22}, s'_{i44}, s'_{i41}; i=1, 2, 4

Assumption (5′) [X exogenous]

Potential response type proportions are the same in both X strata, {X=0} and {X=1}. This implies that $r'_{ijk}/Pr(X=0) = s'_{ijk}/Pr(X=1) = q'_{ijk}$. Use q'_{ijk} in place of r'_{ijk} and s'_{ijk} as the LP variables to be solved for objective function optimization. Replace the parameters p′(y, z, x) by p′(y, z| x). The eight equality constraints are then identical to those under assumption (5) in Section A.1.1, with q′ replacing q and p′ replacing p

The objective functions are modified by replacing r′ with q′ and deleting the s′ terms.

Acknowledgments

This work was supported by Contract R01-HD-39746 from the National Institute of Child Health and Human Development. We are grateful: to Alexander A Balke for use of his symbolic LP solutions computer program; to Dr. M. Alan Brookhart, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Harvard Medical School, for access to the data from his article used as an example in Section 4.1 of this paper; and to Dr. E. Jane Costello, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, for access to the data from the Great Smoke Mountains Study used as an example in Section 4.2 of this paper.

References

Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Am Stat Assoc 1996;91:444–55.

Balke, AA. Dept of Computer Science. University of California: Los Angeles; 1995. Probabilistic Counterfactuals: Semantics, Computation, and Applications. Technical Report R-242, Cognitive Systems Laboratory. Ph.D. dissertation

Balke AA, Pearl J. Bounds on treatment effects from studies with imperfect compliance. J Am Stat Assoc 1997;92:1171–6.

Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. Epidemiology 2006;17:268–75. [PubMed: 16617275]

Cole SR, Hernán MA. Fallibility in estimating direct effects. Int J Epidem 2002;31:163–5.

Copas JB. Randomization models for the matched and unmatched 2×2 tables. Biometrika 1973;60:467–76.

Costello EJ, Compton SN, Keeler G, Angold A. Relationships between poverty and psychopathology: a natural experiment. JAMA 2003;290:2023–9. [PubMed: 14559956]

Flanders WD, Khoury MJ. Indirect assessment of confounding: Graphic description and limits on effect of adjusting for covariates. Epidemiology 1990;1:239–46. [PubMed: 2081259]

Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. Int J Epidemiol 1986;15:433–9. [PubMed: 3771089]

Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? Epidemiology 2006;17:360–72. [PubMed: 16755261]

Kaufman S, Kaufman JS, Maclehose RF, Greenland S, Poole C. Improved estimation of controlled direct effects in the presence of unmeasured confounding of intermediate variables. Stat Med 2005;24:1683–702. [PubMed: 15742358]

Lin DY, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. Biometrics 1998;54:948–63. [PubMed: 9750244]

MacLehose RF, Kaufman S, Kauman JS, Poole C. Bounding causal effects under uncontrolled confounding using counterfactuals. Epidemiology 2005;16:548–55. [PubMed: 15951674]

Manski C. Nonparametric bounds on treatment effects. American Economic Review Papers and Proceedings 1990;80:319–23.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Mattheiss TH. An algorithm for determining irrelevant constraints and all vertices in systems of linear inequalities. Operations Research 1973;21:247–60.

Pearl, J. Causality: Models, Reasoning and Inference. Cambridge, UK: Cambridge University Press; 2000.

Pearl, J. Cognitive Systems Laboratory, Computer Science Department. University of California: Los Angeles; 2001. Direct and indirect effects. Technical Report R-273.

Petersen ML, Sinisi SE, van der Laan MJ. Estimation of direct causal effects. Epidemiology 2006;17:276–84. [PubMed: 16617276]

Robins, JM. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest, L.; Freeman, H.; Mulley, A., editors. Health Service Research Methodology: A Focus on AIDS. Washington, D.C: U.S Public Health Service, National Center for Health Services Research; 1989. p. 113-59.

Robins, JM. Semantics of causal DAG models and the identification of direct and indirect effects. In: Green, P.; Hjort, NL.; Richardson, S., editors. Highly Structured Stochastic Systems. New York: Oxford University Press; 2003. p. 70-82.

Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology 1992;3:143–55. [PubMed: 1576220]

Robins JM, Greenland S. Identification of causal effects using instrumental variables: Comment. J Am Stat Assoc 1996;91:456–8.

Rothman, KJ.; Greenland, S. Basic methods for sensitivity analysis and external adjustment. Chapter 19. In: Rothman, KJ.; Greenland, S., editors. Modern epidemiology. Second edition. Philadelphia: Lippincott, Williams & Wilkins; 1998.

Rubin DB. Formal modes of statistical inference for causal effects. J Stat Plan Inference 1990;25:279–92.

Sommer A, Zeger SL. On estimating efficacy from clinical-trials. Stat Med 1991;10:45–52. [PubMed: 2006355]

Yanagawa T. Case-control studies: Assessing the effect of a confounding factor. Biometrika 1984;71:191–4.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

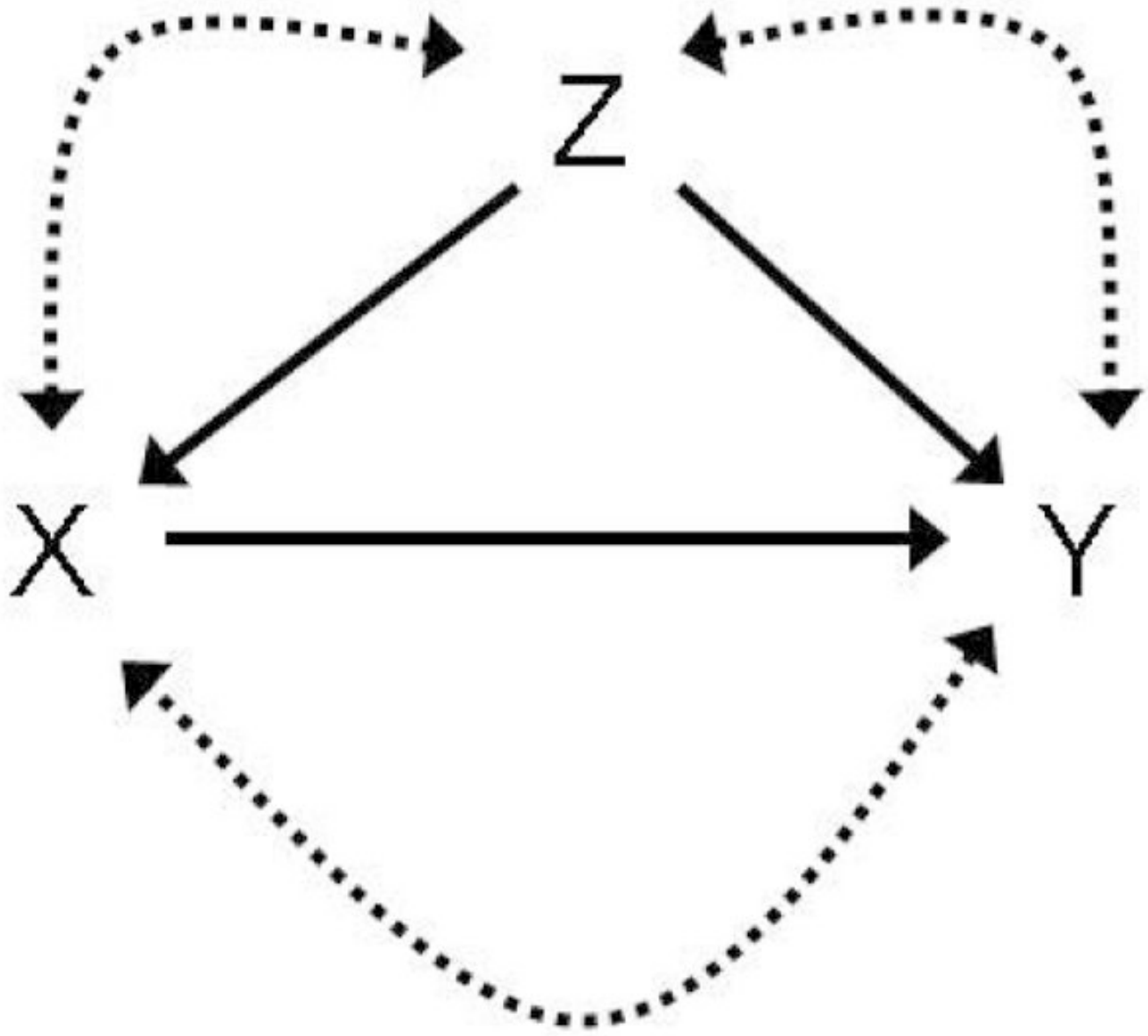

Figure 1.
The Pretreatment Covariate DAG

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

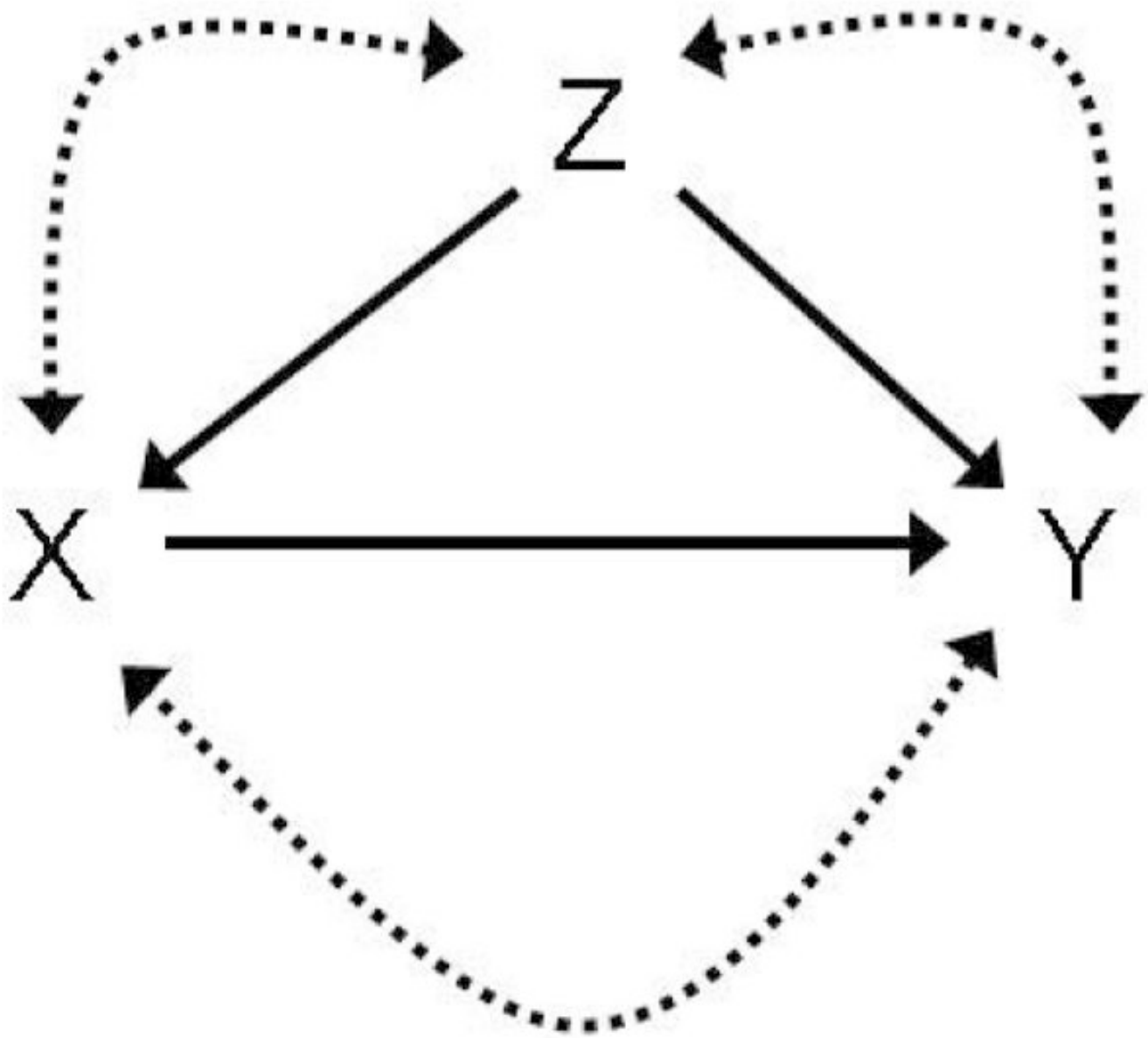

Figure 2.
The Intermediate Covariate DAG

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

TABLE 1

Analytic Bounds on Risk Difference Causal Effects in the Pretreatment-covariate DAG [Z →X →Y, Z →Y] under Various Assumptions

Assumption[1]	Effect Measure[2]	Lower Bound	Upper Bound
None	RD_C	$-\Pr(Y\neq X)$	$\Pr(Y=X)$
	$RD_{C\|Z=z}$	$-\Pr(Y\neq X\|Z=z)$	$\Pr(Y=X\|Z=z)$
	$RD_{C\|SET[Z=z]}$	$-\Pr(Y\neq X)-\Pr(Y=X,Z\neq z)$	$\Pr(Y=X)+\Pr(Y\neq X, Z\neq z)$
{1}	RD_C, $RD_{C\|SET[Z=z]}$	$-\Pr(Y\neq X)$	$\Pr(Y=X)$
	$RD_{C\|Z=z}$	$-\Pr(Y\neq X \|Z=z)$	$\Pr(Y=X\|Z=z)$
{2}	RD_C	0	$\Pr(Y=X)$
	$RD_{C\|Z=z}$	0	$\Pr(Y=X\|Z=z)$
	$RD_{C\|SET[Z=z]}$	0	$\Pr(Y=X)+\Pr(Y\neq X, Z\neq z)$
{3}	RD_C	0	$\Pr(Y=X)$
	$RD_{C\|Z=z}$	0	$\Pr(Y=X\|Z=z)$
	$RD_{C\|SET[Z=z]}$	0	$\Pr(Y=X)+\Pr(Y\neq z, X=z, Z\neq z)$
{4}	All	0	$\min\{\Pr(Y=X\|Z=0), \Pr(Y=X\|Z=1)\}$
{5}	RD_C	$-\Pr(Y\neq X)$	$\Pr(Y=X)$
	$RD_{C\|Z=z}$, $RD_{C\|SET[Z=z]}$	$-\Pr(Y\neq X\|Z=z)$	$\Pr(Y=X\|Z=z)$
{1,5}	All	$\max\{a_1,\ldots, a_8\}$[3]	$\min\{b_1,\ldots, b_8\}$[3]
{2,5}	RD_C	0	$\Pr(Y=X)$
	$RD_{C\|Z=z}$, $RD_{C\|SET[Z=z]}$	0	$\Pr(Y=X\|Z=z)$
{3,5}	RD_C	0	$\min\left\{\begin{array}{l}\Pr(Y=X),\\ 1-p(1,0\mid 0)-p(0,1\mid 1),\\ \Pr(Y=X)+p(1,0,1)-p(1,0,0),\\ \Pr(Y=X\mid Z=1)+p(1,0,1)-p(1,0,0),\end{array}\right\}$
	$RD_{C\|Z=z}$, $RD_{C\|SET[Z=z]}$	0	$\min\left\{\begin{array}{l}\Pr(Y=X\mid Z=z),\\ 1-p(1,0\mid 0)-p(0,1\mid 1)\end{array}\right\}$
{4,5}	All	0	$\min\left\{\begin{array}{l}\Pr(Y=X\mid Z=0),\\ \Pr(Y=X\mid Z=1),\\ \Pr(Y=X\mid Z=0)+p(1,0\mid 1)-p(1,0\mid 0),\\ \Pr(Y=X\mid Z=1)+p(1,0\mid 0)-p(1,0\mid 1)\end{array}\right\}$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Assumption[1]	**Effect Measure**[2]	**Lower Bound**	**Upper Bound**
{1,2}, {1,3}, {1,4}	RD_C, $RD_{C\|SET[Z=z]}$	0	Pr(Y=X)
	$RD_{C\|Z=z}$	0	Pr(Y=X|Z=z)
{1,2,5}, {1,3,5}, {1,4,5}	All	Pr(Y=1|Z=1) − Pr(Y=1|Z=0)	Pr(Y=1|Z=1) − Pr(Y=1|Z=0) + p(1,1| 0) + p(0,0| 1)

[1] Assumptions are defined in Section 2.3

{1} No direct Z →Y effect

{2} Partial monotonicity

{3} Full monotonicity

{4} Full monotonicity and no interaction between Z and X

{5} Exogenous (e.g. randomized)

[2] Effect Measures are defined in Section 2.2

RD_C = Pr(Y=1| SET[X=1]) − Pr(Y=1| SET[X=0])

$RD_{C|Z=z}$ = Pr(Y=1| SET[X=1], Z=z) − Pr(Y=1| SET[X=0], Z=z)

$RD_{C|SET[Z=z]}$ = Pr(Y=1| SET[X=1], SET[Z=z]) − Pr(Y=1| SET[X=0], SET[Z=z])

[3] Balke & Pearl bounds (1997)

[4] p notation:

p(y,x,z) = Pr(Y=y, X=x, Z=z)

p(y,x | z)= Pr(Y=y, X=x | Z=z)

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

TABLE 2

Analytic Bounds on Risk Difference Causal Effects in the Intermediate-covariate DAG [X →Z →Y, X →Y] under Various Assumptions

Assumption[1]	Effect[2]	Lower Bound	Upper Bound
None	Total	−Pr(Y≠X)	Pr(Y=X)
	CDE	−Pr(Y≠X)−Pr(Y=X, Z≠z)	Pr(Y=X)+Pr(Y≠X, Z≠z)
	NDE	−Pr(Y≠X)−Pr(Y=X=1)	Pr(Y=X)+Pr(Y≠X=1)
{1′}	Total	−Pr(Y≠X)	Pr(Y=X)
{2′}	Total	−Pr(Y≠X)	Pr(Y=X)
	CDE	−Pr(Y≠X)−Pr(Y=X=Z≠z)	Pr(Y=X)+Pr(X=z, Y=Z≠z)
	NDE	−Pr(Y≠X)−Pr(Y=X=Z=1)	Pr(Y=X)
{3′}	Total, NDE	0	Pr(Y=X)
	CDE	0	Pr(Y=X)+Pr(X=z, Y=Z≠z)
{4′}	Total, CDE, NDE	0	Pr(Y=X)
{5′}	Total[3]	Pr(Y=1|X=1)−Pr(Y=1|X=0)	Pr(Y=1|X=1)−Pr(Y=1|X=0)
	CDE	−1+Pr(Y=0,Z=z|X=0) + Pr(Y=1,Z=z|X=1)	1−Pr(Y=1,Z=z|X=0)−Pr(Y=0,Z=z|X=1)
	NDE[4]	$\max\left\{\begin{array}{c}-\Pr(Y=1\mid X=0),\\ -1+p'(0,1\mid 0)-p'(1,0\mid 0)+p'(1,1\mid 1),\\ -1+p'(0,0\mid 0)-p'(1,1\mid 0)+p'(1,0\mid 1)\end{array}\right\}$	$\min\left\{\begin{array}{c}\Pr(Y=0\mid X=0),\\ 1+p'(0,1\mid 0)-p'(1,0\mid 0)-p'(0,0\mid 1),\\ 1+p'(0,0\mid 0)-p'(1,1\mid 0)-p'(0,1\mid 1)\end{array}\right\}$
{1′,2′}, {1′,3′}, {1′,4′}	Total	0	Pr(Y=X=Z)
{2′,5′}	CDE	Pr(Y=1|X=1) − Pr(Y=1|X=0) − Pr(Y=Z≠z|X≠z)	Pr(Y=1|X=1) − Pr(Y=1|X=0) + Pr(Y=Z≠z|X=z)
	NDE	$\max\left\{\begin{array}{c}p'(1,0\mid 1)-p'(0,1\mid 1)+p'(0,1\mid 0)-p'(1,0\mid 0),\\ p'(1,0\mid 1)-p'(1,1\mid 0)-p'(1,0\mid 0)\end{array}\right\}$	Pr(Y=1|X=1) − Pr(Y=1|X=0)

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Assumption[1]	Effect[2]	Lower Bound	Upper Bound
{3′,5′}	CDE	$\max\left\{\begin{array}{c}0,\\ \Pr(Y\neq Z=z\mid X\neq z)-\Pr(Y\neq Z=z\mid X=z)\end{array}\right\}$	$\Pr(Y\neq z\mid X\neq z)-\Pr(Y\neq Z=z\mid X=z)$
	NDE	$\max\left\{\begin{array}{c}0,\\ p'(0,1\mid 0)-p'(0,1\mid 1),\\ p'(1,0\mid 1)-p'(1,0\mid 0),\\ p'(0,1\mid 0)-p'(0,1\mid 1)+p'(1,0\mid 1)-p'(1,0\mid 0)\end{array}\right\}$	$\Pr(Y=1\mid X=1)-\Pr(Y=1\mid X=0)$
{4′,5′}	CDE	$\max\left\{\begin{array}{c}0,\\ \Pr(Y\neq z,Z=z\mid X\neq z)-\Pr(Y\neq z,Z=z\mid X=z),\\ \Pr(Y=z,Z\neq z\mid X=z)-\Pr(Y=z,Z\neq z\mid X\neq z),\\ \Pr(Y\neq z,Z=z\mid X\neq z)-\Pr(Y\neq z,Z=z\mid X=z)+\\ \Pr(Y=z,Z\neq z\mid X=z)-\Pr(Y=z,Z\neq z\mid X\neq z)\end{array}\right\}$	$\Pr(Y=1\mid X=1)-\Pr(Y=1\mid X=0)$
	NDE	$\max\left\{\begin{array}{c}0,\\ p'(0,1\mid 0)-p'(0,1\mid 1),\\ p'(1,0\mid 1)-p'(1,0\mid 0),\\ p'(0,1\mid 0)-p'0,1\mid 1)+p'(1,0\mid 1)-p'(1,0\mid 0)\end{array}\right\}$	$\Pr(Y=1\mid X=1)-\Pr(Y=1\mid X=0)$

[1] Assumptions are defined in Section 3.3

{1} No direct X →Y effect

{2} Partial monotonicity

{3} Full monotonicity

{4} Full monotonicity and no interaction between Z and X

{5} Exogenous (e.g. randomized)

[2] Effect Measures are defined in Section 3.1

Total = Total Average Causal Effect

RD_C = Pr(Y=1| SET[X=1]) − Pr(Y=1| SET[X=0])

CDE = Z-Controlled Direct Effect

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

$RD_{C|SET[Z=z]}$ = Pr(Y=1| SET[X=1], SET[Z=z])−Pr(Y=1| SET[X=0], SET[Z=z])

NDE = Natural Direct Effect

$RD_{C|SET[Z=Z(0)]}$ = Pr(Y=1| SET[X=1], SET[Z=Z_0]) −Pr(Y=1| SET[X=0])

[3] Causal effect is completely determined (bound width = 0)

[4] p′ notation:

p′(y,z,x) = Pr(Y=y, Z=z, X=x)

p′(y,z | x)= Pr(Y=y, Z=z | X=x)