



# Using directed acyclic graphs to guide analyses of neighbourhood health effects: an introduction

N L Fleischer, A V Diez Roux

University of Michigan, Ann Arbor, Michigan, USA

Correspondence to:  
Dr A V Diez Roux, Center for Social Epidemiology and Population Health, 3rd Floor SPH Tower, 109 Observatory St, Ann Arbor, MI 48109-2029, USA; adiezrou@umich.edu

Accepted 22 October 2007

## ABSTRACT

**Background:** Directed acyclic graphs, or DAGs, are a useful graphical tool in epidemiologic research that can help identify appropriate analytical strategies in addition to potential unintended consequences of commonly used methods such as conditioning on mediators. The use of DAGs can be particularly informative in the study of the causal effects of social factors on health.

**Methods:** The authors consider four specific scenarios in which DAGs may be useful to neighbourhood health effects researchers: (1) identifying variables that need to be adjusted for in estimating neighbourhood health effects, (2) identifying the unintended consequences of estimating “direct” effects by conditioning on a mediator, (3) using DAGs to understand possible sources and consequences of selection bias in neighbourhood health effects research, and (4) using DAGs to identify the consequences of adjustment for variables affected by prior exposure.

**Conclusions:** The authors present simplified sample DAGs for each scenario and discuss the insights that can be gleaned from the DAGs in each case and the implications these have for analytical approaches.

Directed acyclic graphs, or DAGs, have emerged as a potentially useful tool in epidemiologic research.<sup>1–6</sup> By working through these causal diagrams which graphically encode relationships between variables, epidemiologists can refine their research questions and decide on appropriate analytic plans. The study of the causal effects of social factors on health is one area of epidemiologic inquiry where DAGs may be useful.<sup>7</sup> Investigation of the causal effects of social factors requires dealing with complex causal chains involving multiple inter-related variables ranging from distal antecedents to the more proximal biologic precursors of disease. Glymour<sup>7</sup> has reviewed the fundamentals of DAGs and their potential utility in social epidemiology, but applications to specific research problems in social epidemiology remain rare. Through a series of examples we illustrate the possible use of DAGs in neighbourhood health effects research and the lessons that can be learned from them.

Simplification is a necessary step in science; DAGs are useful in that they make these often unstated simplifications explicit so that their implications can be evaluated. In the spirit of using DAGs as a simplifying tool rather than as an all-encompassing representation of reality, we consider four simplified scenarios in which the use of DAGs can yield important insights in neighbourhood health effects research. In developing these examples we draw heavily on prior work by Hernán and others.<sup>2,7,8</sup> We consider four specific

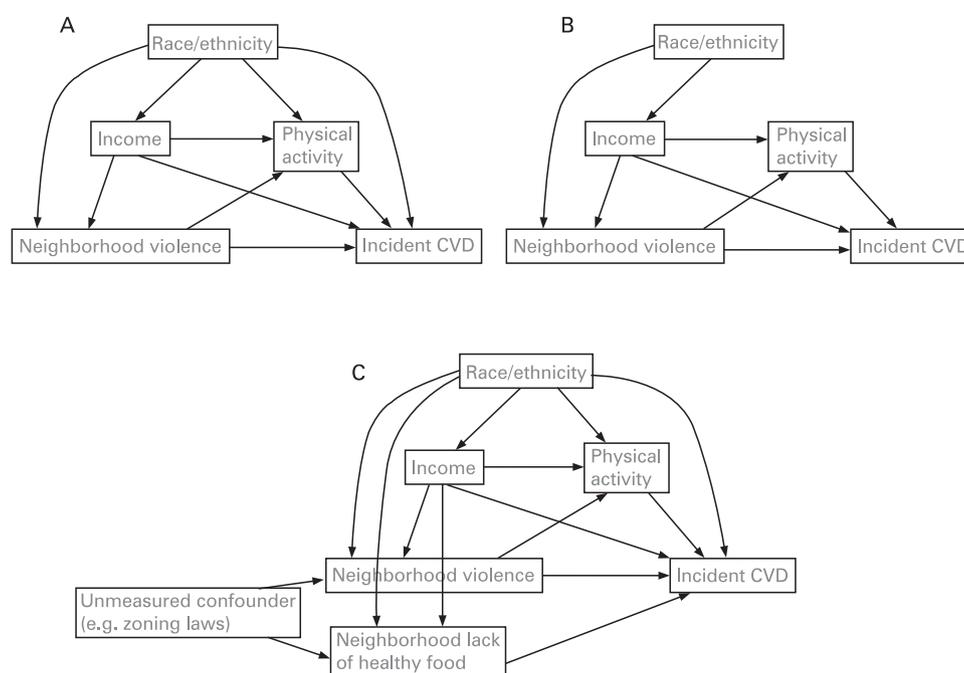
scenarios: (1) identifying variables that need to be adjusted for in estimating neighbourhood health effects, (2) identifying the unintended consequences of estimating “direct” effects by conditioning on a mediator, (3) using DAGs to understand possible sources and consequences of selection bias in neighbourhood health effects research, and (4) using DAGs to identify the consequences of adjustment for variables affected by prior exposure.

Estimating the population causal effect<sup>9–11</sup> requires assigning the exposure (a neighbourhood attribute in this case) to each individual and measuring the outcome in each individual, that is, it requires individual-level data. For this reason, we refer to individual-level exposures to neighbourhood characteristics (an individual-level attribute). Thus, all DAGs we will discuss are individual-level DAGs (in which the units of analysis are individuals and all variables are measured for each individual). In estimating this causal effect, methods that account for non-independence of outcomes within higher level units or over space generally (such as multilevel or spatial models) may be necessary, but these are not discussed further in this paper. For the purposes of simplification we also assume no heterogeneity of individual-level effects across higher level units and no cross-level interactions (ie, the neighbourhood effect is assumed to be homogeneous across levels of individual characteristics). Also for simplicity, we will assume a dichotomous exposure variable. Although the examples we use are based on the types of research questions that are often investigated in the neighbourhood effects research, they are purely hypothetical and are used merely to illustrate the concepts discussed. They do not indicate the presence of empirical support for any particular scenario.

## USING DAGS TO IDENTIFY VARIABLES THAT NEED TO BE ADJUSTED FOR IN ESTIMATING NEIGHBOURHOOD HEALTH EFFECTS

Figure 1A shows a simple DAG illustrating the hypothesised relationship between a neighbourhood characteristic and incidence of cardiovascular disease (CVD). In estimating the causal effect of a neighbourhood characteristic, such as violence, on the development of CVD, we also consider the effects of race/ethnicity, socioeconomic position (measured by income) and behaviour, such as physical activity. Since neighbourhood violence could affect an individual's outdoor activity in his or her local area, physical activity is an intermediate on an indirect causal pathway between the neighbourhood characteristic and CVD. Race/ethnicity

**Figure 1** Using directed acyclic graphs to identify variables that need to be controlled for in estimating neighbourhood health effects.



and income affect physical activity and are also causally associated with CVD through pathways that do not involve physical activity (such as other behaviours or the stress process). Race/ethnicity affects income (eg, through access to education and occupational opportunities).

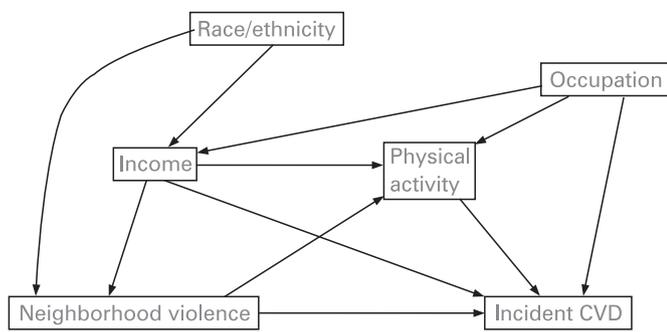
The arrows from race/ethnicity and individual-level income to neighbourhood violence represent the hypothesised causal effect of race/ethnicity or income on an individual's exposure to neighbourhood violence. Although individual-level income or race/ethnicity do not themselves "cause" neighbourhood violence, these individual-level characteristics are causally related to an individual's probability of living in a neighbourhood with a certain level of violence, through processes involving residential choice, constraints and discrimination. An important distinction arises in this context because of the multilevel nature of the research problem. The causes of an individual's exposure to neighbourhood violence include individual-level characteristics that result in a person living in a neighbourhood with high levels of violence (income and race/ethnicity in our example) and causes of the neighbourhood violence itself. The causes of neighbourhood violence are defined at the neighbourhood level (including, for example, emergent properties such as level of cohesion between neighbours and their willingness to intervene for the common good<sup>12</sup>) or at a higher level (such as racial residential segregation at the level of the larger metropolitan area). For simplification, these higher level causes of neighbourhood-level violence are not shown in most of the DAGs we discuss.

In fig 1A, the total effect of the neighbourhood characteristic, violence, on CVD can be estimated by comparing the overall difference in risk between persons whose neighbourhood has the specific characteristic and persons whose neighbourhood does not. However, this simple comparison may be confounded. Although epidemiologists have long used intuition, a priori knowledge and simple rules (such as identifying variables that are associated with the exposure and the outcome) to identify confounders, recent work has illustrated how this approach can sometimes lead to incorrect decisions.<sup>2</sup>

DAGs allow researchers to use relatively simple and systematic graphical criteria to identify the set of variables  $S$  that needs to be controlled for in order to identify the causal effects of interest. Set  $S$  is sufficient to control for confounding if there is no confounding of the neighbourhood violence–CVD risk relationship in any stratum of  $S$ .<sup>2</sup> In DAG terminology,  $S$  is sufficient for adjustment if (1) no variable in  $S$  is a descendent of (or caused by) the exposure, (2) every unblocked backdoor path from exposure to outcome is intercepted by a variable in  $S$  (which blocks the path) and (3) every unblocked path between exposure and outcome induced by adjustment for the variables in  $S$  (as discussed in more detail below) is intercepted by a variable in  $S$ .<sup>2,7</sup> (An unblocked path is a sequence of arrows (regardless of the direction of the arrows) connecting two variables that does not contain a collider. A collider is a variable with two arrows pointing into it. For example, in fig 2, physical activity is a collider on the violence–physical activity–occupation–CVD path; therefore, the path is blocked. A backdoor path is a path that begins with an arrow pointing into the exposure and ends in an arrow pointing into the outcome. For example, in fig 1a, the neighbourhood violence–race/ethnicity–CVD path is an unblocked backdoor path since there is no collider and the path begins with an arrow pointing into violence and ends with an arrow pointing into CVD. A variable is said to be a child of another variable if it is caused by it. A descendent of a variable is a child of a variable or another variable further down a causal path (eg, a child of a child). Additional details on DAG terminology can be found in Pearl<sup>3</sup> and Glymour<sup>7</sup>). In fig 1a, a set  $S$  of covariates whose control may be necessary to eliminate confounding of the total effect of neighbourhood violence on CVD can be identified using the steps below.

#### Steps to determine set $S$ of covariates necessary to control for confounding

1. Delete all arrows emanating from neighbourhood violence (to CVD, physical activity).
2. Now see whether there are any unblocked backdoor paths from exposure to disease.



**Figure 2** Using directed acyclic graphs to identify the unintended consequences of estimating “direct” effects by conditioning on a mediator.

3. There are unblocked backdoor paths from neighbourhood violence via race/ethnicity and income to CVD, so race/ethnicity and income will need to be controlled for.

We can also examine whether a given set (eg, income and race/ethnicity) is sufficient to identify the total effect of neighbourhood violence on CVD by using the graphical algorithm known as the test for minimal sufficiency.<sup>2</sup>

#### Steps to determine whether set *S* of covariates is *minimally sufficient* to control for confounding

1. Delete all arrows emanating from the exposure.
2. Draw undirected arcs (lines without arrows) to connect every pair of variables that share a child that is in *S* or has a descendant in *S* (these are associations generated by the control of *S*, as will be discussed in more detail below).
3. In the new graph generated in 2 see whether there is any unblocked path (ie, sequences of lines connecting variables without a collider) from exposure to disease that does not pass through *S*. These are new paths generated by control for *S*. If these are present the set *S* is not sufficient, and a variable intercepting this path must be added to *S*.

In fig 1A adjustment for income or race/ethnicity alone is insufficient to estimate the total causal effect of neighbourhood violence because both strategies would leave unblocked backdoor paths from CVD to neighbourhood (through race/ethnicity in the case of adjustment for income and through income in the case of adjustment for race/ethnicity). If, however, as shown in fig 1B, race/ethnicity is related to physical activity and CVD only through its effects on income, then adjustment for race/ethnicity would be unnecessary if one has already adjusted for income. In fact, if race/ethnicity is strongly related to neighbourhood violence (eg, because residential segregation leads to a strong association between being a member of a certain ethnic group and living in a violent neighbourhood) then the unnecessary adjustment for race/ethnicity could limit our ability to identify the causal effect of neighbourhood violence on CVD because it may not be possible to reliably separate these two effects in our data. In addition, if the neighbourhood-level construct is measured with error, and race/ethnicity is a better measure of exposure to neighbourhood violence than the available measure of neighbourhood-level violence itself, the (unnecessary) adjustment for race/ethnicity could result in no association between neighbourhood violence and CVD, simply because the effect of neighbourhood is proxied by race/ethnicity owing to measurement error in the available neighbourhood-level exposure variable.

Figure 1C illustrates a scenario in which another neighbourhood-level exposure such as absence of availability of healthy

foods shares a common, unmeasured cause with neighbourhood violence (eg, zoning laws) and is causally related to CVD incidence through a separate mechanism. Under this scenario, adjustment for race/ethnicity and income would still leave an unblocked backdoor path from CVD to neighbourhood violence; hence, availability of healthy foods would have to be controlled for to estimate the total effect of violence of CVD risk.

#### USING DAGS TO IDENTIFY THE UNINTENDED CONSEQUENCES OF ESTIMATING A “DIRECT” EFFECT BY CONDITIONING ON A MEDIATOR

A common question in neighbourhood health effects research pertains to the estimation of the “direct” effect of a neighbourhood characteristic, or, in other words, the effect that operates through processes that do not involve measured mediators. In fig 1B, for example, researchers may be interested in effects of neighbourhood violence on CVD that operate through mechanisms other than physical activity. Traditionally, the direct effect is estimated after controlling for confounders of the total effect plus mediators of the indirect effect. An important insight from DAGs is that in some cases adjusting for a mediator introduces a new source of bias. This situation arises when the relationship between the mediator and the outcome is confounded by a third variable (which is not a confounder of the total effect). In fig 2, for example, occupation is not a confounder of the total effect of neighbourhood violence on incident CVD once income is controlled. However, by statistically controlling for or conditioning on physical activity in order to estimate the direct effect, the investigator induces an association between occupation and the neighbourhood characteristic within strata of physical activity; occupation thus becomes a confounder of the “direct effect”.<sup>1 13</sup> This association is induced because physical activity is a collider on the neighbourhood violence–physical activity–occupation path.

An intuitive explanation of this is that if neighbourhood violence and low occupation are causes of less physical activity, and we know that a given individual has low physical activity but does not live in a violent neighbourhood, we know that that individual is likely to be in a low occupation category (the other cause of low physical activity). In other words, in the absence of living in a violent neighbourhood, it is likely that the other cause of low physical activity (low occupation) is present. Thus, neighbourhood violence and occupation are associated (are not independent) within strata of the physical activity level. Because of this, in order to estimate the direct effect we must control for occupation in addition to income and physical activity (even though occupation is not a confounder of the total effect once income is controlled). If we apply the test for minimal sufficiency to income and physical activity, in this example, we would see that neighbourhood violence and occupation share a descendant (physical activity) in *S*; therefore, adjustment for *S* creates an association between occupation and neighbourhood violence within strata of *S* (see step 2 of the criteria for minimal sufficiency). Therefore, occupation must be added to *S* in order to estimate the direct effect.

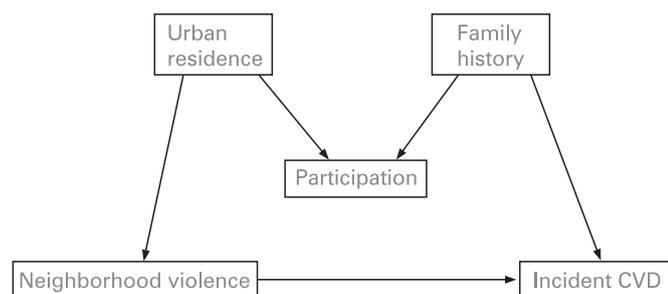
The extent to which controlling for mediators results in under- or overestimates of the direct effect will depend on whether confounders of the mediator–outcomes relationship are present and on the strength and directionality of the confounding.<sup>13</sup> In our example, the consequences of not accounting for occupation when estimating the direct effect depends on the strength and directionality of the associations of occupation with physical activity and CVD risk. Even if

unmeasured confounders of the mediator–outcome association are present, it is possible that the consequences of their omission in estimating the direct effect are trivial compared with other issues such as measurement error in confounders of the total effect (eg, income in fig 2) or measurement error in the mediators (eg, physical activity in fig 2).<sup>14</sup>

### USING DAGS TO UNDERSTAND POSSIBLE SOURCES AND CONSEQUENCES OF SELECTION BIAS IN NEIGHBOURHOOD HEALTH EFFECTS RESEARCH

Hernán *et al*<sup>8</sup> have used DAGs to show how selection bias can arise when researchers condition on a common effect of exposure (or a cause of exposure) and outcome (or a cause of the outcome). Figure 3 shows an example in which participation in a study of neighbourhood health effects on CVD is affected by urban residence (with urban residents being more likely to participate) and family history (with persons with a family history of CVD being more likely to participate). Urban residence is a cause of exposure (because urban areas are more likely to have higher levels of neighbourhood violence), and family history is a cause of CVD (through genetic or other shared family factors). In the full population neither urban residence nor family history are confounders of the association between neighbourhood violence and CVD. However, when conditioning on participation in the study, family history and urban residence become associated, creating a non-causal link between CVD and neighbourhood violence. This arises because, if urban residence and family history are both causes of participation and a participant does not have a family history, he or she is more likely to live in an urban area. Thus, urban residence and family history may be associated among participants even if they are unassociated in the full population.

Analogous to the situation discussed for physical activity in fig 2, conditioning on participation induces an association between exposure and outcome; this problem is commonly referred to in epidemiology as selection bias. Judging the directionality and strength of the induced association can be complex.<sup>15</sup> It is also important to note that, in addition to a spurious association due to selection bias, any observed association between neighbourhood violence and CVD among participants could also arise if exposure to violence causes CVD only among urban residents with a family history. Thus, the observed association among participants would result from a combination of selection bias and true causal association among urban residents with a family history. The relative contribution of these two processes to the observed association cannot be inferred from this simple DAG.

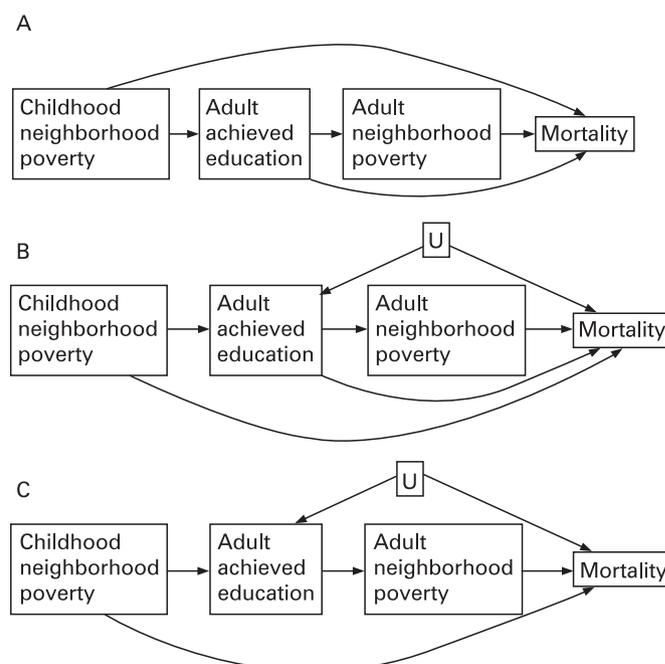


**Figure 3** Using directed acyclic graphs to understand possible sources and consequences of selection bias in neighbourhood health effects research.

### USING DAGS TO IDENTIFY THE CONSEQUENCES OF ADJUSTMENT FOR VARIABLES AFFECTED BY PRIOR EXPOSURE

The processes through which neighbourhoods affect health may involve the effects of cumulative exposure to adverse neighbourhood conditions. In order to estimate the cumulative effects it may be necessary to adjust for variables which are affected by prior exposure to neighbourhood conditions, but adjustment for these variables using the common approach of regression adjustment (or the equivalent stratification) can result in bias.<sup>8</sup> Special methods (such as marginal structural models or structural nested models<sup>5 15 16</sup>) may be necessary to correctly estimate the cumulative effects of interest. DAGs are helpful in identifying these situations.

A simple example is shown in fig 4A. Neighbourhood conditions early in childhood affect achieved education in adulthood, and achieved education in adulthood affects subsequent residential location (and therefore exposure to neighbourhood poverty) later in life and mortality. Therefore, education is simultaneously a confounder and a mediator of the cumulative effect of neighbourhood poverty on mortality because it mediates the effects of early life neighbourhood conditions but confounds the effects of adult neighbourhood conditions. Under these circumstances, conditioning on education will result in an unbiased estimate of the “direct” effect of lifecourse neighbourhood poverty on mortality (assuming the DAG is correct and no omitted confounders of the education/mortality relationship are present). However, the total effect of neighbourhood poverty on mortality cannot be estimated using standard regression techniques because the effect of adult neighbourhood poverty is confounded by education, but education mediates the portion of the cumulative effect that results from childhood neighbourhood poverty. Hence, neither the unadjusted nor the education-adjusted association of cumulative neighbourhood poverty with mortality provides an unbiased estimate of the total effect of cumulative neighbourhood poverty on mortality.



**Figure 4** Using directed acyclic graphs to identify the consequences of adjustment for variables affected by prior exposure. This figure is modelled in part on a similar figure in the paper by Hernán *et al*.<sup>8</sup>

Figure 4B shows a slightly more complex scenario in which there is an unmeasured confounder of the relationship between education and mortality. In this case, neither the direct nor the total effect of lifecourse neighbourhood poverty can be estimated using standard approaches.  $U$  confounds the relationship between education and mortality. Any approach that stratifies on adult education will create a spurious association between childhood poverty and  $U$  (because  $U$  and childhood poverty are causes of education). But, at the same time, adult education is a confounder of the relationship between adult neighbourhood poverty and mortality. The “direct” effect of lifecourse poverty cannot be estimated without bias because controlling for education (which is a mediator but also a collider on the childhood poverty–education– $U$ –mortality path) creates a spurious association between childhood neighbourhood poverty and mortality. The total effect of lifecourse poverty cannot be estimated without bias because education is a confounder of the adult portion of the neighbourhood poverty lifecourse exposure, but controlling for it introduces bias.

A similar problem is present in fig 4C: even if education were not a mediator of the effect of childhood poverty on mortality, adjustment for education would be necessary to correctly estimate the adult portion of the cumulative effect but would introduce bias in the estimate of the childhood effect. Hence, the total cumulative effect cannot be estimated without bias using standard approaches. However, adjustment for education is appropriate and introduces no bias if the intent is to estimate only the effect of adult neighbourhood poverty. In the situations illustrated in fig 4A–C the magnitude of the bias will result from the relative importance of the confounding of the effect of adult neighbourhood that is eliminated, and the bias in the effect of neighbourhood poverty that is created when adult education is controlled for.<sup>8</sup> In allowing investigators to identify situations analogous to fig 4A–C, DAGs are helpful in identifying the need for alternative analytical strategies that are not based on stratification or conditioning on covariates.<sup>1–5</sup>

## CONCLUSION

We have used simple examples to illustrate the insights that can be gleaned from DAGs in the investigation of neighbourhood health effects. We focused on the use of DAGs to study how exposure to neighbourhood characteristics is related to individual-level health outcomes, that is, the contribution of exposure to neighbourhood characteristics to variation between individuals in health. For this reason the DAGs we discuss are formulated as individual-level DAGs. It is also possible to ask questions about the causes of neighbourhood-to-neighbourhood variation in neighbourhood-level outcomes, and it would be plausible to construct DAGs to answer neighbourhood-level questions as well. Just as it may be important to consider neighbourhood-level factors in understanding between-individual variation, it may be necessary to consider individual-level factors in understanding between-neighbourhood variation. Combining both types of questions in a single DAG may be possible but is beyond the scope of this paper.

DAGs cannot encode parametric assumptions, strengths of associations, nor statistical interactions. In addition, as noted by Glymour,<sup>7</sup> it is important to go beyond the use of DAGs to highlight possible biases and develop methods to place bounds on the amount of bias likely to be present under different assumptions. For example, data can be simulated based on DAGs, and then analysed to understand the magnitude of potential biases that can result from different analyses.<sup>7</sup> DAGs cannot be easily used to evaluate the impact of reciprocal

## What this study adds

Although DAGs have received increasing attention in epidemiology, specific applications to social epidemiology remain rare. This paper highlights the possible utility of DAGs in neighbourhood health effects research.

## Policy implications

Improving the ability to identify causal effects of neighbourhoods on health may lead to more effective policies for disease prevention.

relations between features of neighbourhoods and the residents that live in them. Other approaches more suited to describing these types of relationships (including complex systems approaches such as agent-based models<sup>17</sup>) may be useful complements. Improving the rigour of observational research of neighbourhood health effects is likely to benefit from many complementary strategies and methods. DAGs are one of several tools in this arsenal.

**Acknowledgements:** The authors would like to thank Katherine J Hoggatt for her helpful comments on a previous version.

**Funding:** This work was supported by R01 HL071759 from the National Heart, Lung, and Blood Institute.

**Competing interests:** None.

## REFERENCES

1. Cole SR, Hernan MA. Fallibility in estimating direct effects. *Int J Epidemiol* 2002;**31**:163–5.
2. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology (Cambridge, Mass)* 1999;**10**:37–48.
3. Pearl J. *Causality: models, reasoning, and inference*. Cambridge: Cambridge University Press, 2000.
4. Robins JM. Data, design, and background knowledge in etiologic inference. *Epidemiology (Cambridge, Mass)* 2001;**12**:313–20.
5. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology (Cambridge, Mass)* 2000;**11**:550–60.
6. Hernandez-Diaz S, Schisterman EF, Hernan MA. The birth weight “paradox” uncovered? *Am J Epidemiol* 2006;**164**:1115–20.
7. Glymour MM. Using causal diagrams to understand common problems in social epidemiology. In: Oakes JM, Kaufman JS, eds. *Methods in social epidemiology*. San Francisco, CA: Jossey-Bass, 2006:393–428.
8. Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology (Cambridge, Mass)* 2004;**15**:615–25.
9. Hernan MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health* 2004;**58**:265–71.
10. Subramanian SV, Glymour MM, Kawachi I. Identifying causal ecologic effects on health: a methodological assessment. In: Galea S, ed. *Macrosocial determinants of population health*. New York: Springer Media, 2007.
11. Winship C, Morgan SL. The estimation of causal effects from observational data. *Annu Rev Sociol* 1999;**25**:659–706.
12. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science* 1997;**277**:918–24.
13. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology (Cambridge, Mass)* 2003;**14**:300–6.
14. Blakely T. Commentary: estimating direct and indirect effects—fallible in theory, but in the real world? *Int J Epidemiol* 2002;**31**:166–7.
15. Robins J. The control of confounding by intermediate variables. *Stat Med* 1989;**8**:679–701.
16. Robins JM. Testing and estimation of direct effects by reparameterizing directed acyclic graphs with structural nested models. In: Glymour CN, Cooper GF, eds. *Computation, causation, and discovery*. Menlo Park, CA: AAAI Press; MIT Press, 1999:349–405.
17. Miller JH, Page SE. *Complex adaptive systems: an introduction to computational models of social life*. Princeton Studies in Complexity. Princeton, NJ: Princeton University Press, 2007.