

**Word count**

**Abstract:** 233

**Text:** 3065

## **Using Causal Diagrams to Improve the Design and Interpretation of Medical Research**

Mahyar Etminan PharmD, MSc<sup>1</sup>, Gary Collins PhD<sup>2</sup>, Mohammad Ali Mansournia MD, MPH, PhD<sup>3</sup>

<sup>1</sup>Department of Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, Canada

<sup>2</sup>Centre for Statistics in Medicine, Nuffield Department of Orthopedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

<sup>3</sup>Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Corresponding Author:

Mohammad Ali Mansournia MD, MPH, PhD

Associate Professor of Epidemiology

Tehran University of Medical Sciences

Tel: +98-21-88989123; Fax: +98-21-88989127

Email: [Mansournia\\_m@sina.tums.ac.ir](mailto:Mansournia_m@sina.tums.ac.ir), [mansournia\\_ma@yahoo.com](mailto:mansournia_ma@yahoo.com)

**Key words:** Causal directed acyclic graphs, colliders, confounding, selection bias, detection bias, over adjustment bias

**Funding:** We received no funding for this work

**Conflicting of interest:** All authors report no conflicts for this work

### **Abstract**

Causal directed acyclic graphs (cDAGs) have become popular tools for researchers to better examine biases related to causal questions. DAGs comprise a series of arrows connecting nodes that represent variables and in doing so can demonstrate the causal relation between different variables. cDAGs can provide researchers with a blueprint of the exposure and

outcome relationship and the other variables that play a role in that causal question. cDAGs can be helpful in the design and interpretation of observational studies in pulmonary, critical care, sleep and cardiovascular medicine. cDAGs can help clinicians and researchers to better identify the structure of different biases that can affect the validity of observational studies.

Most of the available literature on cDAGs and their function use language that might be unfamiliar to clinicians. In this paper, we explain cDAG terminology and the principles behind how they work. We use cDAGs and clinical examples that are mostly focused in the area of pulmonary medicine to describe the structure of confounding, selection bias, overadjustment bias and detection bias. We then apply these principles to a more complex published case study on the use of statins and COPD mortality. We also introduce readers to other resources for a more in-depth discussion of causal inference principles.

## **Introduction**

Medical research often attempts to ascertain the predictors of a certain outcome or whether a treatment causes a certain outcome. Randomized controlled trials (RCTs) are considered the gold standard for establishing causation. As RCT data are not available for most causal questions, many research studies that address causal questions use observational designs.

At times results from observational studies can confuse the effect of interest with other variables' effects, leading to an association that is not causal<sup>1</sup>. It would be helpful for clinicians and researchers to be able to visualize the structure of biases in a clinical study to

ensure that the study design and analysis considers these kinds of issues. By visualizing these biases researchers can better identify and control for them.

In recent years there has been an explosion of observational studies in the areas of pulmonary, critical care, sleep and cardiovascular medicine using Big Data. This has led to the identification of different types of biases that might affect the validity of these studies<sup>2</sup>. Causal directed acyclic graphs (cDAGs) are a relatively new framework for visualizing the variables that play a role in causal questions<sup>3,4</sup>. Several reviews and tutorials have been published on the use of cDAGs intended for clinicians and clinician researchers<sup>5-8</sup>. A review by Lederer<sup>8</sup> introduces general concepts including confounding and selection bias for pulmonary medicine researchers. Here we extend these concepts to include other pertinent terminologies and principles of causal inference that we believe researchers should be familiar with. We also attempt to link principles of causal inference to clinical examples which we believe is critical to the understanding of these concepts for the readers.

In this review we will expand on the existing literature and provide a more detailed review of how cDAGs can identify four common types of biases in epidemiological studies: *confounding, selection, overadjustment, and detection bias* using (for the most part) examples from the pulmonary, critical care, sleep and cardiovascular medicine literature. We will also use a recently published research study on the effect of statin use and COPD mortality as a case study and demonstrate, by drawing a cDAG, how to identify and control for potential biases<sup>9</sup>.

## **Directed acyclic graphs**

Directed acyclic graphs (DAGs) comprise a series of arrows connecting nodes that represent variables. Any series of adjacent arrows in the graph, regardless of direction, forms a path. The series must be acyclic, meaning it cannot form a feedback loop in which one variable causes itself.

DAGs are based on mathematical graph theory and have historically been used in computer science, particularly in the creation of computer algorithms. They are increasingly used to address causation in areas of science such as epidemiology and medicine<sup>10</sup>. DAGs that show a causal effect between two or more variables are called causal DAGs (cDAGs). Table 1 defines common terms related to cDAGs. We will now demonstrate the application of cDAGs in identifying different types of biases using clinical examples.

## **Identifying bias structures in epidemiologic studies using cDAGs**

### *Confounding bias*

Confounding is perhaps the most prevalent type of bias that can affect the validity of epidemiological studies. Consider an observational cohort that uses a large health claims database to examine the risk of new COPD exacerbations with current users of a new long acting beta<sub>2</sub>adrenergic receptor agonist (LABA). Suppose the new LABA had been previously shown to reduce COPD exacerbations in a large randomized trial but in the observational cohort study investigators observe the opposite effect (a harmful effect with the drug). One potential reason for this paradoxical effect is confounding by indication.

Figure 1Sa shows the variables at play in this cohort study. In this figure, arrows represent a direct effect from one variable to another. There is an arrow or *causal path* from LABA

use to COPD exacerbations as this is the causal question we want to examine. This path can also be called a *directed path* as it describes a path that starts from LABA use and ends on COPD exacerbations. Thus, all causal paths are directed paths.

COPD severity is an *ancestor* of COPD exacerbations while COPD exacerbations would be a *descendent* of COPD severity. There is an arrow from COPD severity to LABA use (patients with more severe COPD are more likely to receive LABAs) and from COPD severity to COPD exacerbations because severe COPD can lead to COPD exacerbations.

The path LABA ← COPD severity → COPD exacerbation is an *undirected path* because the sequence of arrows is not directional from LABA use to COPD exacerbation. This undirected path through COPD severity is a *biasing path*. The path on which COPD severity lies is also referred to as a *confounding path* as it is a path that includes a common cause of LABA use and COPD exacerbations. COPD severity introduces confounding bias<sup>7,11</sup> because it is a common cause of COPD exacerbations and LABA use. This type of confounding is referred to as *confounding by indication* (also referred to as confounding by disease severity or channeling bias) because it arises due to more patients on LABAs with more severe COPD developing COPD exacerbation than those taking LABAs with less severe COPD. Biasing paths such as a confounding path need to be blocked in order to ascertain an unbiased measure of the causal effect (Table 2).

Now that we have acknowledged that COPD severity is a confounder, we can attempt to block the path that it creates to prevent this bias from being transmitted. The most common approach for blocking this path is to statistically adjust for COPD severity. Approaches include stratifying patients by COPD severity, matching LABA users and nonusers on COPD severity and restricting or excluding patients on a proxy for this variable. All these methods are examples of conditioning. Once we have conditioned on COPD severity the only causal path remaining in the cDAG is the path from LABA use to COPD exacerbations.

### *Selection bias*

Selection bias is another type of bias that can affect validity of epidemiologic studies. Unlike confounding, selection bias is created by researchers at the study design stage. It occurs when a study conditions on a variable that is a common effect of both the exposure and outcome<sup>7,10</sup>.

Consider a researcher who wants to examine the risk of epilepsy with oral fluoroquinolones, as these drugs have recently been linked to adverse events in the central nervous system by the Food and Drug Administration<sup>12</sup>. Figure 1Sb is a cDAG depicting this scenario.

There are two causal paths from fluoroquinolone use to epilepsy, one a directed path and the other an undirected path that could lead to bias. Two arrows originate from pneumonia and go to fluoroquinolone use and hospitalization, as pneumonia causes both outcomes. Similarly, two arrows originate from electrolyte imbalance and go to epilepsy and hospitalization. As pneumonia and electrolyte imbalance both cause hospitalization, their effects are said to *collide* on hospitalization and hospitalization is called a *collider* on this path<sup>7,13</sup>. As the collider sits on the undirected path between the exposure (fluoroquinolone use) and outcome (epilepsy), it prevents the transmission of any bias through this *blocked path*. The presence of the collider means that this undirected path is not a biasing path. However, this blocked path can be opened upon conditioning on the variable hospitalization (Table 2S).

In our example, imagine a researcher who is not aware of the structure of this cDAG and is curious to know if the risk of epilepsy with fluoroquinolones differs among those who are hospitalized compared with those who are not. They decide to restrict (a form of conditioning) the data to only hospitalized patients. This approach will now open the path that was previously blocked by hospitalization. That is, initially there was no flow of information from fluoroquinolone use to epilepsy due to the blocked path by the variable ‘hospitalization’ and hence no possibility for transmission of bias. Upon conditioning by hospitalization this previously blocked path is now open creating the path: fluoroquinolones ← pneumonia → hospitalization ← electrolyte imbalance → epilepsy. This new path will introduce a biased, non-casual association referred to as a selection bias or *collider bias*<sup>7,13</sup>. Thus, hospitalization should not have been conditioned on. A remedy to this situation would be to block the transmission of this bias by adjusting for pneumonia and/or electrolyte imbalance.

### *Overadjustment bias*

Investigators are often careful to control for all potential confounding variables, usually using statistical adjustment. Peer reviewers and journal editors also want to be convinced that a submitted study has adjusted for all potential confounding variables. Adjusting for variables that should not be adjusted for leads to *overadjustment bias*<sup>14</sup>. Overadjustment bias usually occurs when an intermediate variable (also called a *mediator*<sup>14</sup>) is adjusted for. Mediators are variables that alter the effect of an exposure on the outcome through a specific mechanism.

For example, antipsychotics can increase the risk of diabetes<sup>15</sup>. Consider an investigator who wants to look at the total effect of antipsychotic use on the risk of diabetes in a cohort study. Figure 1Sc shows two causal paths, a direct path from antipsychotics to diabetes and an indirect path from antipsychotics to diabetes through body mass index (BMI). The investigator might choose to stratify the antipsychotic users by their BMI, categorizing this value into a binary variable, >30 or <30. Conditioning on an intermediate variable will block the flow of information from antipsychotic use to diabetes through an indirect path via BMI (Table 2). The resulting association will be biased due to overadjustment bias.

Intermediate or mediator variables should not be conditioned on when the study question assesses an exposure’s overall effect on an outcome. More complex methodologies are needed if an investigator wants to examine how much of the total effect is derived from the intermediate variable<sup>14</sup>.

### *Detection bias*

Measurement error, also called misclassification of the outcome or exposure, can also affect the validity of epidemiological studies. Detection bias is a type of outcome misclassification in which a group of patients are more likely to be diagnosed with a certain condition than another group because of a third variable.

For example, recent studies have linked the use of thiazide diuretics to skin cancer<sup>16,17</sup>. This increase in risk might be driven by detection bias, as shown by the cDAG in Figure 1Sd. There is an arrow from thiazides to skin changes (an adverse event with thiazide diuretics) and an arrow from skin changes to diagnosed skin cancer. There are also arrows from all skin cancer (diagnosed and undiagnosed) to diagnosed skin cancer and skin changes. The cDAG makes it clear that there is no true effect (arrow) from thiazide diuretics to skin cancer. Notice that although skin changes might appear to be a mediator, it actually

is not as we know that thiazides do not cause skin cancer through skin changes. Skin changes is simply an adverse event of thiazides that can lead to a higher number of diagnosed skin cancers. The spurious association is due to detection bias: thiazide users are more likely to go to their doctor with skin change related issues than non-users, which increases the chance of their skin cancer being spotted, regardless of whether it is related to the thiazide.

We can theoretically control for the detection bias by conditioning on or adjusting for skin changes, blocking the pathway thiazides → skin changes → diagnosed skin cancer. However, the variable ‘all skin cancer’ also has an effect on skin changes. Skin changes is a collider on the remaining undirected path: thiazides → skin changes ← all skin cancer → diagnosed skin cancer. Although conditioning on skin changes removes detection bias, it introduces collider bias. In large database studies, one must often assess the trade-off between accepting detection bias or collider bias. If a clinical trial design had been used, this problem could have been mitigated by using better screening for skin cancer for both thiazide users and non-users.

### **Drawing a cDAG and addressing bias: a case study**

We now present a case study to illustrate how to generate a cDAG and use it to identify biasing paths. A group of researchers investigated the effect of statin medications on lung-related mortality in patients with chronic obstructive pulmonary disease (COPD)<sup>9</sup>. Some evidence suggested that statins may lower mortality through their anti-inflammatory properties<sup>18</sup>. They examined this question with a cohort study using COPD patients from a population health database. To control for survival bias (immortal time bias) the investigators started follow up for statin users and non users one year after cohort entry. Several potential confounders were controlled for in the Cox regression model including age, gender, comorbidity, number of physician and hospital visits, geographic location, and income level.

Figure 2Sa shows the cDAG for this study. The dotted arrow from statins to total lung mortality shows that this is the causal question under investigation. The variables that were controlled for are indicated with a box. They were adjusted because they lie on biasing paths. SP represents survival one year after the first statin prescription, which is a condition created by the investigators in the study design. COPD indicates that the cohort was restricted to COPD patients.

All shared causes (causes shared by statin use and lung mortality) must be included in a cDAG. We have therefore added the variables heart disease and smoking to Figure 2Sa. However, to avoid overcrowding the cDAG and aid visualization of the biasing paths, we have omitted gender and age from Figure 2Sa. All variables are shown in a complete version of the cDAG in Appendix Figure 2Sb.

The aim is to keep all causal paths from statin use to lung-related mortality open and block all biasing paths which are open undirected paths between the exposure and outcome. For this question, we need to block biasing paths that include a common cause of statins and mortality by adjusting for confounders. We must also avoid adjusting for colliders or descendants of colliders in other undirected paths, which could introduce collider bias<sup>7,13</sup> (Table 2Sa).

Income level and comorbidity are common causes of statin use and lung-related mortality. Age, number of hospitalizations, number of prescriptions (to prevent overcrowding of the cDAG we did not include the variables number of hospitalizations and number of prescriptions), gender, number of physician visits (MD visits), and

comorbidity are on confounding paths with statin use and mortality. These confounders were correctly adjusted for by the investigators.

Heart disease, smoking (assuming smoking occurred before heart disease), and COPD severity are also confounders of statin use and lung-related mortality. Although smoking and COPD severity were not available in the study database, they are shown in the cDAG as they are also on a path that includes a common cause (income level). Heart disease was measured in the database but not adjusted in the study. The cDAG suggests that it should have been adjusted for because it causes both lung-related mortality and statin use. The path  $\text{statins} \leftarrow \text{heart disease} \leftarrow \text{smoking} \rightarrow \text{lung-related mortality}$  is a confounding path. Either smoking or heart disease should have also been adjusted for.

Two additional biasing paths have been created by the conditioning steps in the study design. Restricting the cohort to COPD patients created the biasing path:  $\text{statin} \leftarrow \text{heart disease} \leftarrow \text{smoking} \rightarrow \text{COPD} \leftarrow \text{income level} \rightarrow \text{lung-related mortality}$ . The collider bias created by conditioning on COPD can be blocked by conditioning on income level or smoking. As income level is already adjusted for, collider bias is averted.

Only patients who had been taking statins for one year were included in the study. This cohort restriction created the biasing path:  $\text{statin} \rightarrow \text{SP} \leftarrow \text{COPD severity} \rightarrow \text{lung-related mortality}$ . This collider bias could not be adjusted for in the analysis because COPD severity is usually not available in most population-based databases.

In summary, the variables adjusted for by the investigators included: age, gender, number of physician visits (MD visits), comorbidity, income level, number of hospitalizations and number of prescriptions received. Based on our cDAG variables heart disease and COPD severity should have also been adjusted for. Heart disease could have been adjusted by the investigators as this data was available to them. COPD severity was an unmeasured confounder and could not have been adjusted for. The cDAG also tells us that two collider biases were introduced: 1) COPD which was controlled for as a result of adjustment for income level, and 2) variable SP which could not be controlled for since it was introduced at the design stage of the study.

### **Causal diagram resources for clinicians**

Those interested in learning more about the basic principles behind causal diagrams and accessing more detailed discussions of identifying and controlling for confounders and colliders are directed towards recent reviews and tutorials aimed at clinicians<sup>5,8</sup>. Harvard University offers an excellent introductory online course on casual inference with worked examples for those looking for a more interactive experience<sup>9</sup>.

Several resources are available for drawing causal diagrams. Simple cDAGs can be drawn using Microsoft® PowerPoint®. Dagitty (<http://www.dagitty.net/>) is a freely available software program that has been specifically built for drawing causal diagrams.

### **Limitations of cDAGs**

Like any methodological or statistical technique, causal diagrams have limitations. They do not convey information about important aspects of causal relations, such as the magnitude of the bias. cDAGs might imply that several sets of variables are sufficient for bias adjustment but offer no further guidance on which to use. The absence of key variables in a cDAG might itself lead to bias, which may be mitigated by researchers proposing

different cDAGs to gain a better sense of the effects of different causal assumptions. cDAGs are insensitive to biases that result from small sample sizes, such as bias due to chance confounding<sup>20</sup> or sparse data bias<sup>21</sup>. In practice, one cannot be certain of the temporality between variables as they may be measured simultaneously. Similarly, if a casual relationship between two variables is misrepresented, the accuracy of the cDAG will be compromised. Although cDAGs are useful for qualitatively conceptualizing causality, advanced quantitative causal models should be used to estimate causal effects<sup>2,22-24</sup>.

### **Summary and conclusion**

Judea Pearl, who pioneered the adaptation of cDAGs from quantitative sciences to medical research, said, “*Causal analysis without graphs is like medicine without anatomy.*” Although cDAGs should not be considered a panacea for the proper conduct of high-quality research, they are a useful tool for contending with confounding and selection use. We hope that this primer will encourage researchers to make use of cDAGs and the resources available to support their use, while also highlighting the limitations of these tools.

### **Contributor and guarantor information**

Drs. Etminan and Mansournia proposed the idea and drafted the manuscript. Dr. Collins reviewed the manuscript and contributed to its content. Dr. Etminan is the guarantor and accepts full responsibility for the content. Dr. Etminan attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

### **Acknowledgement**

We thank Mohit Sodhi, MSc for his invaluable assistance in preparing the figures and proofreading the manuscript. We acknowledge Dr. Jennifer A. de Beyer of the Centre for Statistics in Medicine, University of Oxford, for English language editing.

## References:

1. Mansournia MA, Etminan M, Danaei G, et al. Time varying confounding in observational research. *BMJ* 2017;359:j4587.
2. Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008;167(4):492-9.
3. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol* 2013;42(3):860-9.
4. Mansournia, MA, Higgins JPT, Sterne JA, et al. Biases in Randomized Trials: A Conversation Between Trialists and Epidemiologists. *Epidemiology* 2017;28:54-59.
5. Williamson EJ, Aitken Z, Lawrie J, et al. Causal diagrams. *Respirology* 2014;19:303-311.
6. Stovitz SD, Shrier I. Causal inference for clinicians. *BMJ Evid Based Med*. Published Online First: 14 February 2019.
7. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology* 2008 8:70.
8. Lederer DJ, Bell SC, Branson RD, et al. Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. *Ann Am Thorac Soc* 2019;16(1):22-28.
9. Raymakers AJN, Sadatsafavi M, Sin DD et al. Impact of Statin Drug Use on All-Cause Mortality in Patients With COPD: A Population-Based Cohort Study. *Chest* 2017;152(3):486-493.
10. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10(1):37-48.
11. Hernán MA, Robins JM (2019). *Causal Inference*. Boca Raton: Chapman & Hall/CRC, forthcoming. Page 82-94.
12. <https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-fda-updates-warnings-oral-and-injectable-fluoroquinolone-antibiotics>. Accessed June 2, 2019

13. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15(5):615-625.
14. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 2009;20(4):488-95.
15. Musil R, Obermeier M, Russ P et al. Weight gain and antipsychotics: a drug safety review. *Expert Opin Drug Saf* 2015;14(1):73-96.
16. Pottegård A, Pedersen SA, Schmidt SAJ et al. Association of Hydrochlorothiazide Use and Risk of Malignant Melanoma. *JAMA Intern Med* 2018;178(8):1120-1122.
17. Pedersen SA, Gaist D, Schmidt SAJ et al. Hydrochlorothiazide use and risk of nonmelanoma skin cancer: A nationwide case-control study from Denmark. *J Am Acad Dermatol* 2018;78(4):673-681.
18. Mroz RM, Lisowski P, Tycinska A et al. Anti-inflammatory effects of atorvastatin treatment in chronic obstructive pulmonary disease. A controlled pilot study. *J Physiol Pharmacol* 2015 66(1):111-28.
19. <https://online-learning.harvard.edu/course/causal-diagrams-draw-your-assumptions-your-conclusions> (Accessed July 23, 2019).
20. Greenland S, Mansournia MA. Limitations of individual causal models, causal graphs, and ignorability assumptions, as illustrated by random confounding and design unfaithfulness. *Eur J Epidemiol* 201;10:1101-10.
21. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. *BMJ* 2016;352:i1981.
22. Shakiba M, Mansournia MA, Salari A, et al. Accounting for Time-Varying Confounding in the Relationship Between Obesity and Coronary Heart Disease: Analysis With G-Estimation: The ARIC Study. *Am J Epidemiol* 2018;187(6):1319-1326.
23. Almasi-Hashiani A, Nedjat S, Mansournia MA. Causal Methods for Observational Research: A Primer. *Arch Iran Med* 2018;21(4):164-169.
24. Mokhayeri Y, Hashemi-Nazari SS, Khodakarim S et al. Effects of Hypothetical Interventions on Ischemic Stroke Using Parametric G-Formula. *Stroke* 2019;50(11):3286-3288.

Table 1. Common definitions used in causal directed acyclic graphs

<b>Term</b>	<b>Definition</b>
Arrow	Shows direct causal effect
Box	A box around a variable means that variable is conditioned upon
Path	A series of adjacent arrows regardless of their directions
Ancestor/Descendent	In a directed path from A to B, A is an ancestor of B and B is a descendant of A
Directed path	A path in which all of the arrows lie tail to head (pointing in the same direction)
Undirected path	A path in which some of the arrows lie tail to head and others head to tail (i.e. not all of the arrows point in the same direction)
Collider	A variable that has two arrowheads converging on it within a path
Collider bias	A bias created when a collider variable is conditioned on. Also known as selection bias
Blocked path	A path that contains a non-collider that has been conditioned on or a path that contains a collider, where neither the collider nor its descendant have been conditioned on
Causal path	A directed path from the exposure to the outcome
Biasing path	An open undirected path between the exposure and outcome
Confounding path	An undirected path that contains a common cause of the exposure and outcome
Confounding	Bias created by a confounding path
Confounder	A variable which is on a confounding path and upon conditioning blocks that path. A confounder can include but is not required to be common causes of exposure and outcome variables
Mediator	A variable that indirectly mediates the effect of the exposure on the outcome



Figure 1Sa. Example of confounding by indication. The effect of long acting beta:adrenergic receptor agonist (LABA) on COPD exacerbations can be confounded by COPD severity if LABAs are prescribed to patients with more severe COPD who are already at a higher risk of developing exacerbations.

Figure 1Sb. Example of collider or selection bias. The effect of fluroquinolones on the risk of epilepsy will be biased if the data are restricted to those who are hospitalized. This restriction creates a selection bias that will spuriously associate fluoroquinolones with the risk of epilepsy

Figure 1Sc. Example of overadjustment bias. The causal effect of antipsychotic use on the risk of diabetes will be biased if body mass index, a mediator on the path, is adjusted for

Figure 1Sd. Example of detection bias. An otherwise absent causal effect between thiazide diuretics and skin cancer can be spuriously observed if thiazide diuretic users are more likely to be diagnosed with skin cancer than non-users, because they visit physicians more often due to skin rashes secondary to thiazide use