The Alan Turing Institute

LEEDS Institute for Data Analytics

UNIVERSITY OF LEEDS

Have DAGs fulfilled their promise?

Dr Peter WG Tennant

University of Leeds, Leeds UK

Alan Turing Institute, London, UK

@PWGTennant

WCE

WORLD CONGRESS OF EPIDEMIOLOGY 2024

CAUSAL THINKING LTD

- I am a director of a company that offers causal inference training and research consultancy services
- I therefore potentially benefit from any activity that promotes the need for, or benefits of, causal inference methods

HAVE DAGS FULFILLED THEIR PROMISE?

Greenland, Pearl, & Robins. *Epidemiology*. 1999

Causal Diagrams for Epidemiologic Research

Sander Greenland,[1] Judea Pearl,[2] and James M. Robins[3]

Causal diagrams have a long history of informal use and, more recently, have undergone formal development for applications in expert systems and robotics. We provide an introduction to these developments and their use in epidemiologic research. Causal diagrams can provide a starting point for identifying variables that must be measured and controlled to obtain unconfounded effect estimates. They also provide a method for critical evaluation of traditional epidemiologic criteria for confounding. In particular, they reveal certain heretofore unnoticed shortcomings of those criteria when used in considering multiple potential confounders. We show how to modify the traditional criteria to correct those shortcomings. (Epidemiology 1999;10:37–48)

Keywords: bias, causation, confounding, epidemiologic methods, graphical methods, observational studies.

Summarization of causal links via graphs or diagrams has long been used as an informal aid to causal analysis. Causal graphs in the form of path diagrams are an integral component of path analysis[1] and structural equations modeling.[2] In more recent times, the theory of directed acyclic graphs (DAGs) has been extended to application in expert-systems research.[3,4] In these applications, there is a pressing need for valid formal rules that allow an automated system or robot to deduce correctly the presence or absence of causal links given correct background information and new data. The outgrowth of this research has been the development of a formal theory for evaluating causal effects using the language of causal diagrams.[5,6] Unlike path analysis and structural-equations modeling, this theory does not require parametric assumptions such as linearity.

The theory of causal graphs is equivalent to the G-computation theory of Robins.[7–9] It has a benefit, however, of providing a compact graphical as well as algebraic formulation of assumptions and results, which may be easier for the general reader to comprehend. In addition, it provides a novel perspective on traditional epidemiologic criteria for confounder identification. This perspective reveals how traditional criteria can be inadequate when multiple confounders must be considered simultaneously. We describe the modifications required of traditional criteria that enable their valid extension to situations involving multiple confounders.

We here provide a brief introduction to the theory of causal diagrams based on DAGs.[5,6] We pay special attention to its relation to nongraphical epidemiologic treatments of confounding.[10–13] We show how diagrams can serve as a visual yet logically rigorous aid for summarizing assumptions about a problem and for identifying variables that must be measured and controlled to obtain unconfounded effect estimates given those assumptions. Thus, use of such graphs can aid in planning of data collection and analysis, in communication of results, and in avoiding subtle pitfalls of confounder selection.

Except where noted otherwise, the present paper will deal only with relations among variables in a given source population; that is, we will deal only with structural (systematic) relations among the underlying variables of interest, so that issues of measurement error and random variation will not arise. We will also not present proofs of results, but we will give references in which proofs can be found.

A Rationale for Graphs

Any deduction about a causal relation must start from some set of assumptions, which we call the analysis model.[14] For example, such a deduction may assume that uncontrolled confounding is negligible; this assumption usually corresponds to a set of assumptions that various uncontrolled factors have negligible associations with

- “*Causal diagrams can provide a starting point for* ***identifying variables that must be measured and controlled to obtain unconfounded effect estimates***”

- “*Such graphs can aid in:*
  - ***Planning of data collection and analysis…***
  - ***Communication of results…***
  - ***Avoiding subtle pitfalls of confounder selection***”

Shrier and Platt. *BMC Medical Research Methodology*. 2008

Correspondence

Open Access

Reducing bias through directed acyclic graphs

Ian Shrier*[1] and Robert W Platt[2]

Address: [1]Centre for Clinical Epidemiology and Community Studies, SMBD-Jewish General Hospital, McGill University, Montreal, Canada and [2]Department of Epidemiology and Biostatistics, McGill University, Montreal, Canada

Email: Ian Shrier* - ian.shrier@mcgill.ca; Robert W Platt - robert.platt@mcgill.ca

* Corresponding author

Published: 30 October 2008

BMC Medical Research Methodology 2008, 8:70 doi:10.1186/1471-2288-8-70

This article is available from: http://www.biomedcentral.com/1471-2288/8/70

Received: 11 March 2008

Accepted: 30 October 2008

Abstract

Background: The objective of most biomedical research is to determine an unbiased estimate of effect for an exposure on an outcome, i.e. to make causal inferences about the exposure. Recent developments in epidemiology have shown that traditional methods of identifying confounding and adjusting for confounding may be inadequate.

Discussion: The traditional methods of adjusting for "potential confounders" may introduce conditional associations and bias rather than minimize it. Although previous published articles have discussed the role of the causal directed acyclic graph approach (DAGs) with respect to confounding, many clinical problems require complicated DAGs and therefore investigators may continue to use traditional practices because they do not have the tools necessary to properly use the DAG approach. The purpose of this manuscript is to demonstrate a simple 6-step approach to the use of DAGs, and also to explain why the method works from a conceptual point of view.

Summary: Using the simple 6-step DAG approach to confounding and selection bias discussed is likely to reduce the degree of bias for the effect estimate in the chosen statistical model.

- "The DAG approach can be used to ***help choose which covariates should be included in traditional statistical approaches in order to minimize the magnitude of the bias in the estimate produced***"
- "…***to help understand whether bias is potentially reduced or increased when conditioning on covariates***"

THE PROMISES

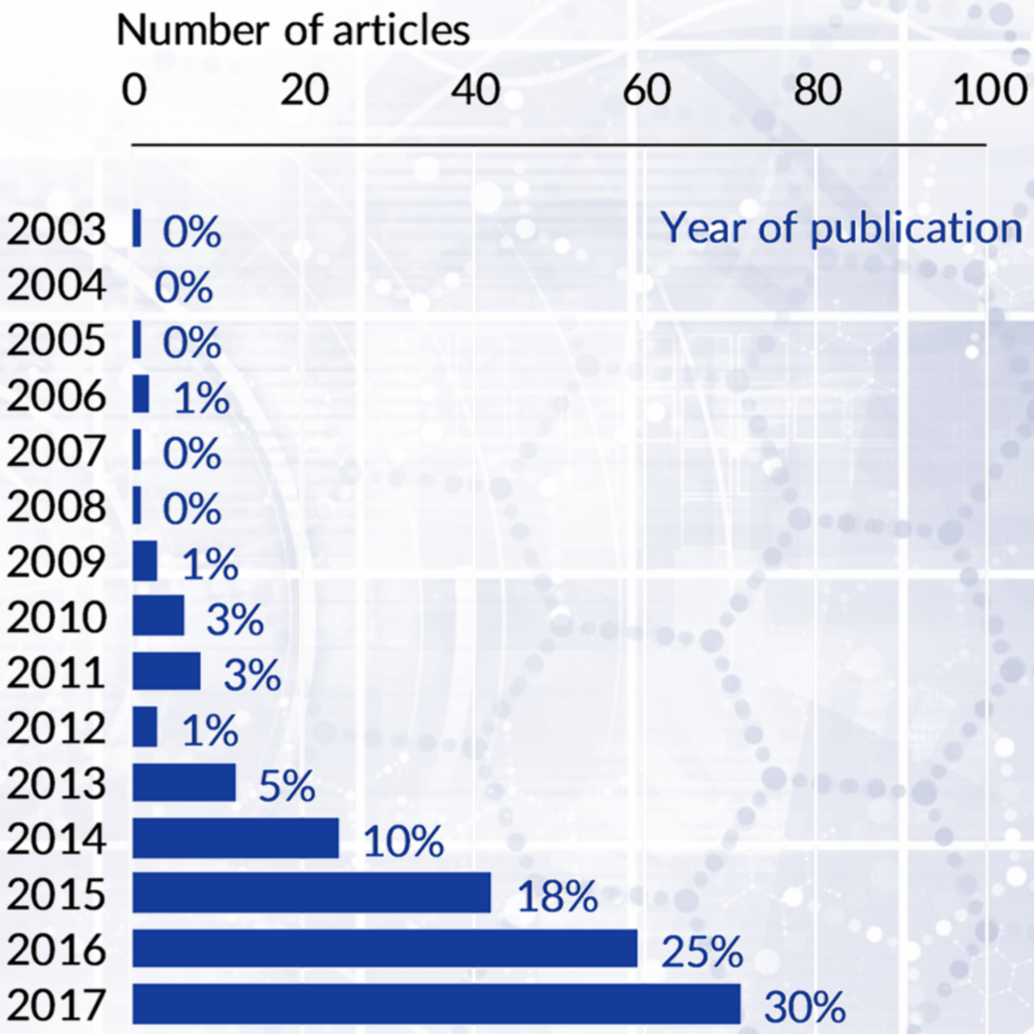
- **1) Helping identify variables that require measuring and adjusting to estimate a causal effect**
  - **Including in planning data collection and analysis**

- **2) Communicating results**
  - **Including communicating assumptions**

- **3) Helping to understand and avoid different types of bias**
  - **Including understanding when conditioning *introduces* bias**

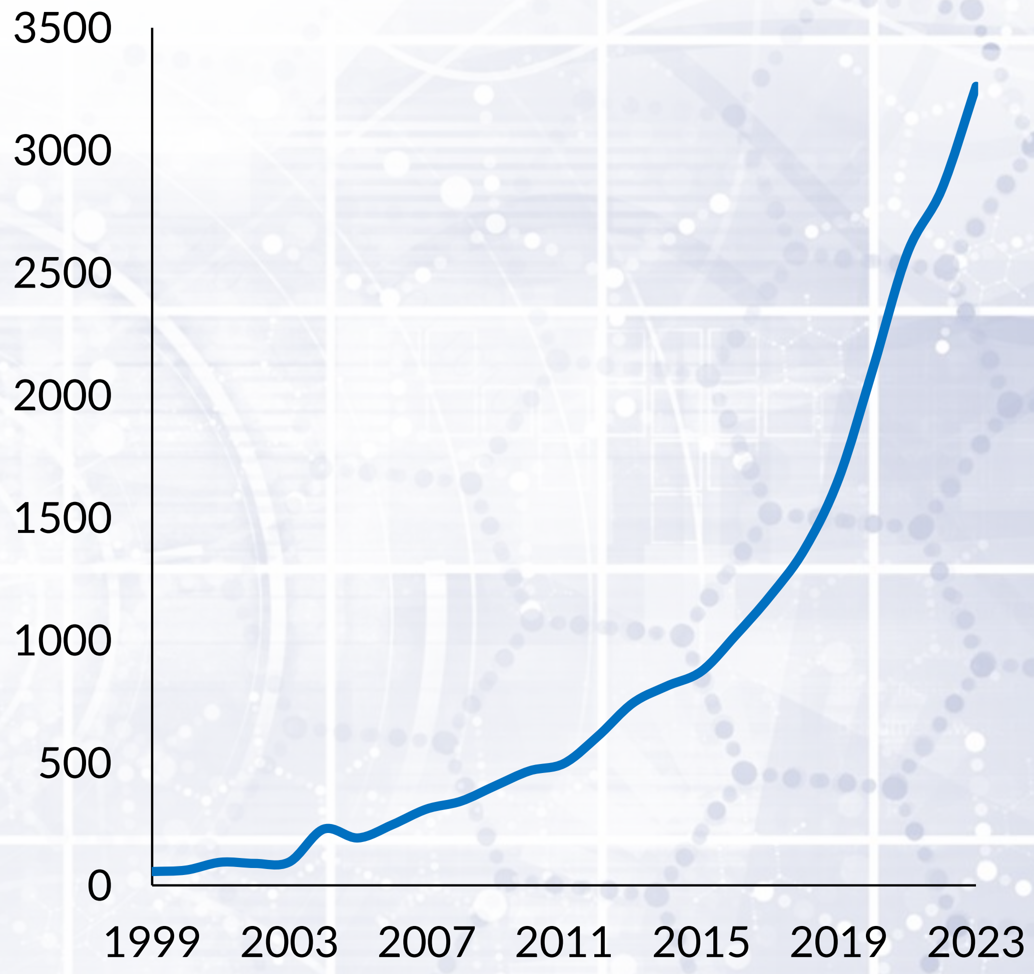
PROMISE 1: HELPING IDENTIFY VARIABLES

- DAGs are a very popular way to identify variables for adjustment
- 2021 review of DAGs identified hundreds of studies using DAGs to identify variables for adjustment



From Tennant et al 2021, IJE

- DAGs are a very popular way to identify variables for adjustment
- 2021 review of DAGs identified hundreds of studies using DAGs to identify variables for adjustment
- Total numbers likely to be far greater!



Google Scholar: “Directed acyclic graph” AND (“Medicine” OR “Health”)

An Overview of Causal Directed Acyclic Graphs for Substance Abuse Researchers

Directed acyclic graphs in perioperative observational research—A systematic review and critique against best practice recommendations

The dawn of directed acyclic graphs in primary care research and education

An introduction to causal inference for pharmacometricians

Causal Diagram Techniques for Urologic Oncology Research

Directed acyclic graphs: a tool for causal studies in paediatrics

Using Directed Acyclic Graphs for Investigating Causal Paths for Cardiovascular Disease

Causal inference in suicide research: When you should (and should not!) control for extraneous variables

An introduction to directed acyclic graphs in trauma research.

Causal inference in drug discovery and development

An Introduction to Causal Diagrams for Anesthesiology Research

Directed Acyclic Graphs in Surgical Research

Directed Acyclic Graphs for Oral Disease Research

Directed acyclic graphs: An under-utilized tool for child maltreatment research

How to implement directed acyclic graphs to reduce bias in addiction research

Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data

A biologist's guide to model selection and causal inference

- DAGs are now strongly encouraged by many causal inference guidelines

NICE real-world evidence framework

Developers should outline their assumptions about the causal relationships between interventions, covariates and outcomes of interest. Ideally, this would be done using causal diagrams known as directed acyclic graphs (Shrier and Platt 2008).

PERSPECTIVE

SPECIAL SECTION

Control of Confounding and Reporting of Results in Causal Inference Studies

Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals

Key Principle #1: Causal inference requires careful consideration of confounding

- Preferred variable selection methods
  1. Historical confounder definition with purposeful variable selection
  2. Causal models using directed acyclic graphs
- Variable selection methods that do not adequately control for confounding
  3. *P* value– or model-based methods
  4. Methods based on β-coefficient changes
  5. Selection of variables to identify "independent predictors"

Guidelines for Reporting Observational Research in Urology: The Importance of Clear Reference to Causality

Check for updates

a. Describe possible causal pathways in the "Methods" section. Although this can be done formally, for instance, using directed acyclic graphs, it is also reasonable to describe causal pathways using ordinary language in the main text.

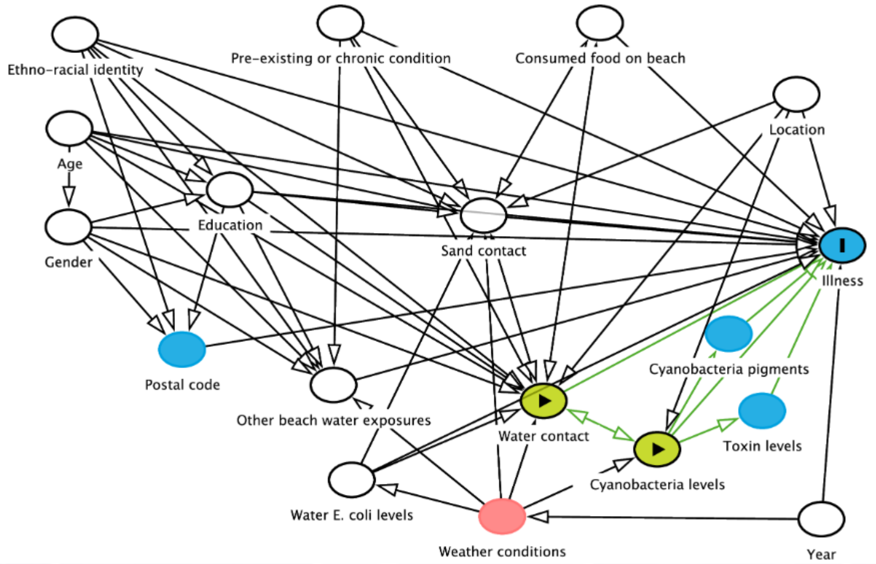
PROMISE 1: HELPING IDENTIFY VARIABLES

- DAGs are increasingly included in protocols to justify data collection and planned modelling strategies

BMJ Open

Burden of recreational water illness due to exposure to cyanobacteria and their toxins in freshwater beaches in Canada: protocol of a prospective cohort study

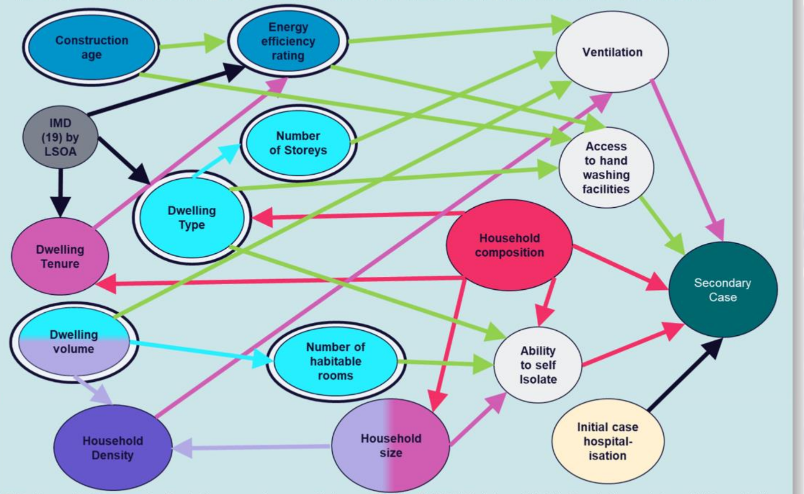
Ian Young,[1] J Johanna Sanchez,[1] Fatih Sekercioglu,[1] Binyam N Desta,[1] Claire Holeton,[2] Dylan Lyng,[3] Victoria Peczulis,[4] Shane Renwick,[5] Teresa Brooks,[6] Jordan Tustin[1]



Development and use of a directed acyclic graph (DAG) for conceptual framework and study protocol development exploring relationships between dwelling characteristics and household transmission of COVID-19 – England, 2020

Hannah Taylor[a,b,c,d,*], Helen Crabbe[b], Clare Humphreys[c], Gavin Dabrera[e], Anna Mavrogianni[f], Neville Q. Verlander[g], Giovanni S. Leonardi[b,h]

Figure 1: Directed Acyclic Graph (DAG) of Environmental Housing Characteristic Exposures and the Outcomes of Becoming a Secondary Case from Household Transmission of SARS-CoV-2



EPC: Energy Performance Certificate (measure of energy efficiency of dwelling). **IMD (19):** Index of Multiple Deprivation rank from 2019 – ranked by
LSOA: Lower Super Output Areas – small areas of population of approximately 1,500 residents.

- **Key benefits:**

1. Encourage us to think explicitly about what variables to collect and control, and make this process much more transparent

List all potential confounding variables

Show your assumptions about potential confounders using a DAG

PROMISE 1: HELPING IDENTIFY VARIABLES

- **Key benefits:**

1. Encourage us to think explicitly about what variables to collect and control, and make this process much more transparent

2. Support wider stakeholder input

Confounders identified by the statistician

Confounders identified by consulting diverse stakeholders

- The visual nature of DAGs makes them accessible even for people with little knowledge of the rules

- From this, people intuitively understand:
  - X occurs before Y
  - X influences Y
- This facilitates *much* wider input into model design than traditional approaches

Reflection on modern methods: constructing directed acyclic graphs (DAGs) with domain experts for health services research

Daniela Rodrigues,[1]* **Noemi Kreif**,[2] **Anna Lawrence-Jones**,[1] **Mauricio Barahona**[3] and **Erik Mayer**[1]

[1]NIHR Imperial Patient Safety Translational Research Centre, Institute of Global Health Innovation, Department of Surgery & Cancer, Imperial College London, London, UK, [2]Centre for Health Economics, University of York, York, UK and [3]Centre for Mathematics of Precision Healthcare, Department of Mathematics, Imperial College London, London, UK

*Corresponding author. NIHR Imperial Patient Safety Translational Research Centre, Institute of Global Health Innovation, Department of Surgery & Cancer, Imperial College London, 10th Floor, Queen Elizabeth the Queen Mother Wing (QEQM), St Mary's Campus, London W2 1NY, UK. E-mail: d.rodrigues@imperial.ac.uk

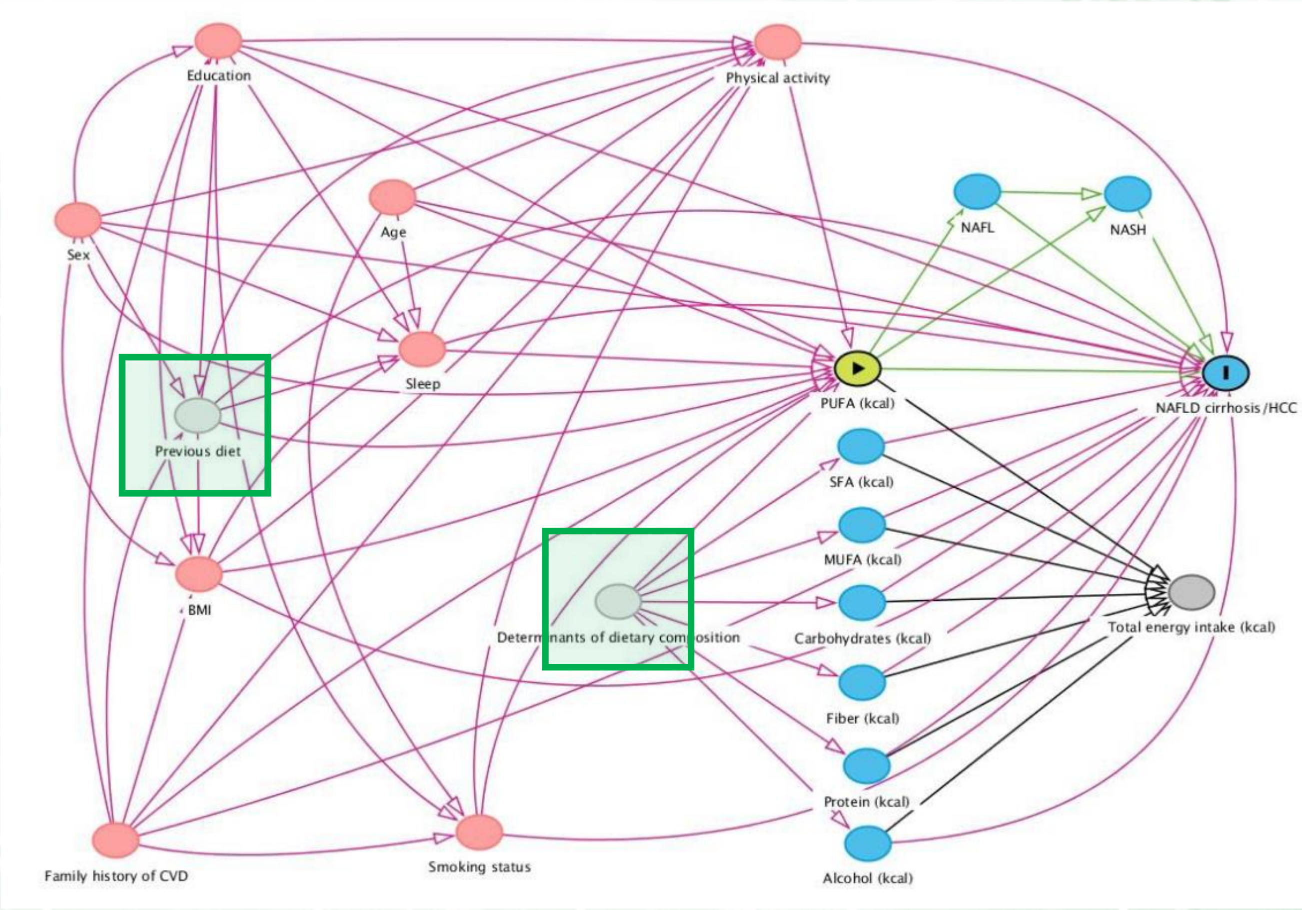
Received 18 August 2021; Editorial decision 22 May 2022; Accepted 7 June 2022

Abstract

Directed acyclic graphs (DAGs) are a useful tool to represent, in a graphical format, researchers' assumptions about the causal structure among variables while providing a rationale for the choice of confounding variables to adjust for. With origins in the field of probabilistic graphical modelling, DAGs are yet to be widely adopted in applied health research, where causal assumptions are frequently made for the purpose of evaluating health services initiatives. In this context, there is still limited practical guidance on how to construct and use DAGs. Some progress has recently been made in terms of building DAGs based on studies from the literature, but an area that has received less attention is how to create DAGs from information provided by domain experts, an approach of particular importance when there is limited published information about the intervention under study. This approach offers the opportunity for findings to be more robust and relevant to patients, carers and the public, and more likely to inform policy and clinical practice. This article draws lessons from a stakeholder workshop involving patients, health care professionals, researchers, commissioners and representatives from industry, whose objective was to draw DAGs for a complex intervention—online consultation, i.e. written exchange between the patient and health care professional using an online system—in the context of the English National Health Service. We provide some initial, practical guidance to those interested in engaging with domain experts to develop DAGs.

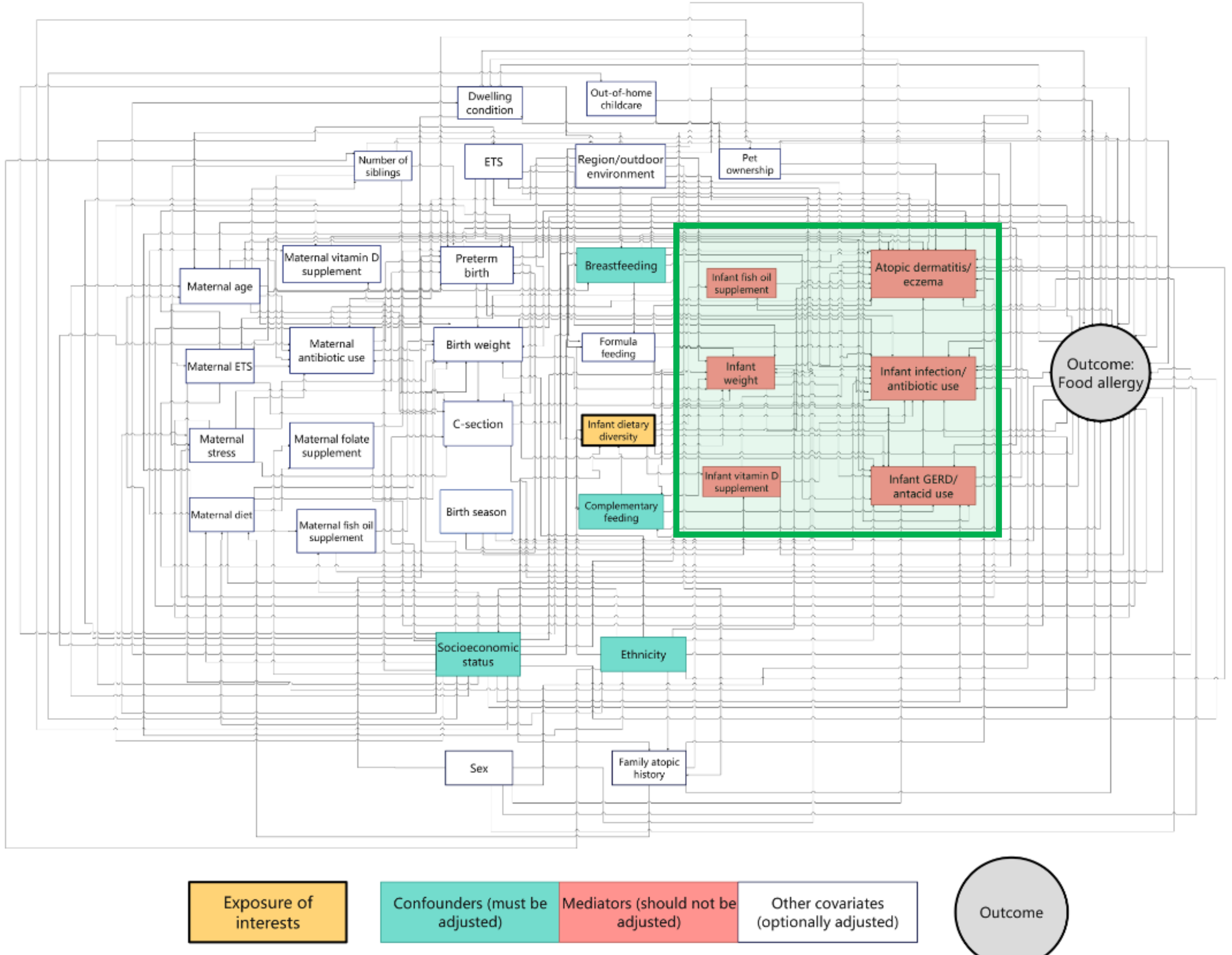
Key words: Causal inference, potential outcomes, directed acyclic graphs, policy evaluation, health services research

- DAGs offer a huge improvement in transparency making data generating assumptions explicit and open to scrutiny
- Great for highlighting residual sources of bias (e.g. unobserved confounding)



Source: Friden et al 2024 *Am J Clin Nutr*

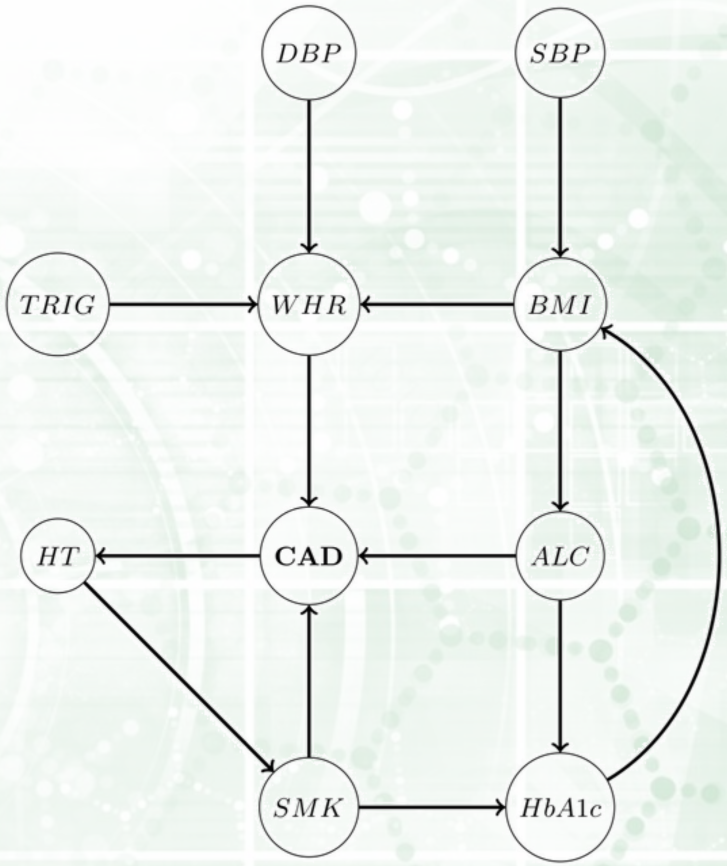
- DAGs offer a huge improvement in transparency making data generating assumptions explicit and open to scrutiny
- Great for highlighting residual sources of bias (e.g. unobserved confounding)
- Great for explaining 'what's going on' and why you *shouldn't* adjust for certain variables



Source: Peng et al 2024 *Allergy*

PROMISE 2: COMMUNICATING RESULTS

- DAGs offer a huge improvement in transparency making data generating assumptions explicit and open to scrutiny
- Great for highlighting residual sources of bias (e.g. unobserved confounding)
- Great for explaining 'what's going on' and why you *shouldn't* adjust for certain variables
- Great for highlighting implausible assumptions!



Assumes:
Blood pressure causes BMI
Cardiovascular disease causes height
BMI causes alcohol consumption

- DAGs can help *everyone* understand key statistical concepts
- E.g. Just 2-hours training plus a DAG-drawing tool can substantially improve probabilistic reasoning among lay people

Widening Access to Bayesian Problem Solving

Nicole Cruz[1], Saoirse Connor Desai[2], Stephen Dewitt[3], Ulrike Hahn[1], David Lagnado[3], Alice Liefgreen[3], Kirsty Phillips[1], Toby Pilditch[3] and Marko Tešić[1]*

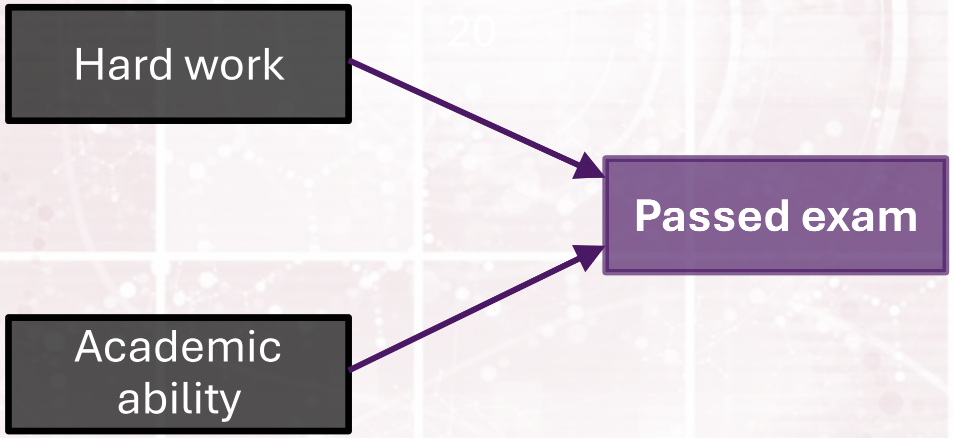
[1] Department of Psychological Sciences, Birkbeck, University of London, London, United Kingdom, [2] Department of Psychology, City, University of London, London, United Kingdom, [3] Department of Experimental Psychology, University College London, London, United Kingdom

Bayesian reasoning and decision making is widely considered normative because it minimizes prediction error in a coherent way. However, it is often difficult to apply Bayesian principles to complex real world problems, which typically have many unknowns and interconnected variables. Bayesian network modeling techniques make it possible to model such problems and obtain precise predictions about the causal impact that changing the value of one variable may have on the values of other variables connected to it. But Bayesian modeling is itself complex, and has until now remained largely inaccessible to lay people. In a large scale lab experiment, we provide proof of principle that a Bayesian network modeling tool, adapted to provide basic training and guidance on the modeling process to beginners without requiring knowledge of the mathematical machinery working behind the scenes, significantly helps lay people find normative Bayesian solutions to complex problems, compared to generic training on probabilistic reasoning. We discuss the implications of this finding for the use of Bayesian network software tools in applied contexts such as security, medical, forensic, economic or environmental decision making.

Keywords: Bayesian networks, assistive software technology, reasoning, decision making, probabilistic

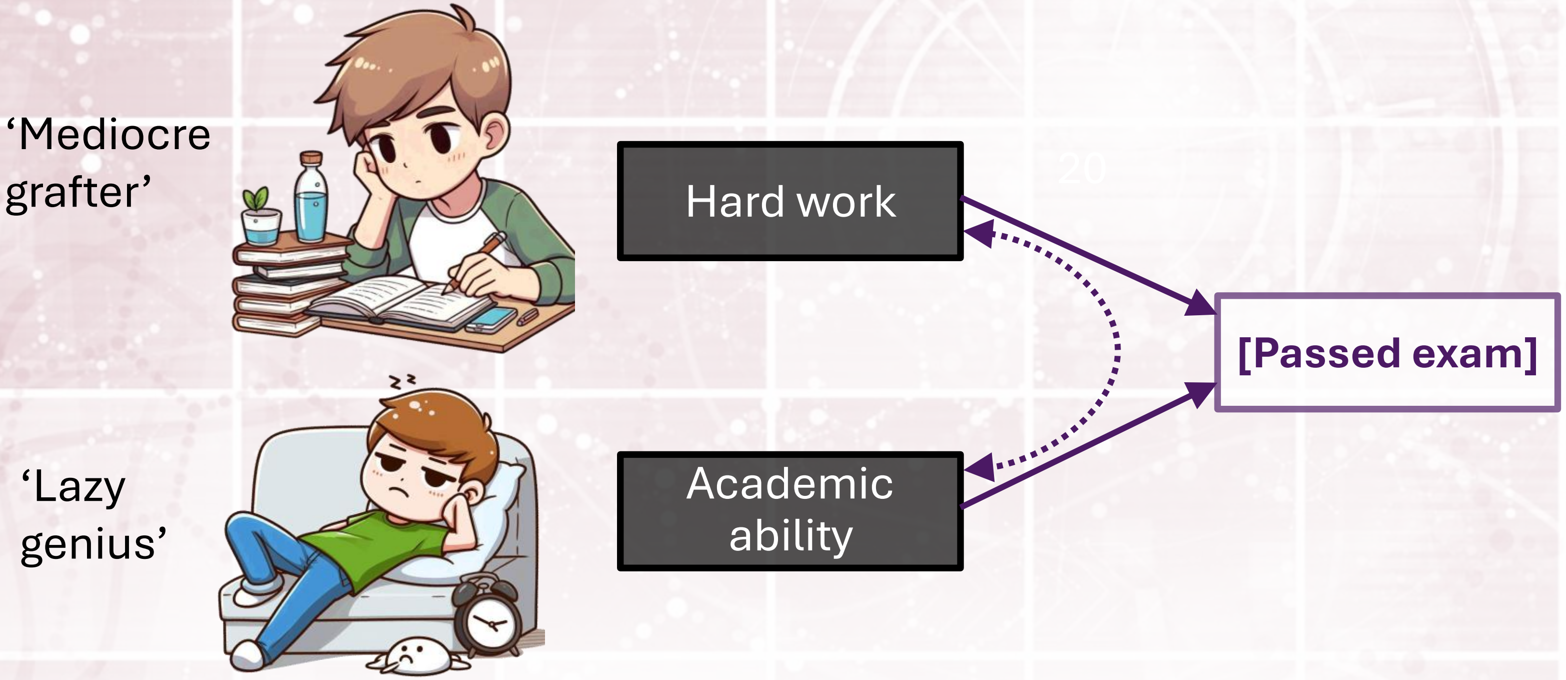
- Among epidemiologists, DAGs have revolutionised our understanding of conditional dependencies – i.e. non-causal associations due to inadvertent or inappropriate conditioning

- E.g. the ‘lazy genius’ and the ‘mediocre grafter’ stereotype



- Among epidemiologists, DAGs have revolutionised our understanding of conditional dependencies – i.e. non-causal associations due to inadvertent or inappropriate conditioning

- E.g. the 'lazy genius' and the 'mediocre grafter' stereotype



- This has been provided many groundbreaking insights!

1. A step change in our understanding of selection bias

A Structural Approach to Selection Bias

Miguel A. Hernán, Sonia Hernández-Díaz,† and James M. Robins**

Collider scope: when selection bias can substantially influence observed associations

Marcus R Munafò,[1,2]* Kate Tilling,[1,3] Amy E Taylor,[1,2] David M Evans,[1,4] and George Davey Smith[1,3]

Toward a Clearer Definition of Selection Bias When Estimating Causal Effects

Haidong Lu,[a] Stephen R. Cole,[b] Chanelle J. Howe,[c] and Daniel Westreich[b]

- This has been provided many groundbreaking insights!

1. A step change in our understanding of selection bias

2. A step change in our understanding of overadjustment bias (and the Table 2 Fallacy)

Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies

Enrique F. Schisterman,[a] Stephen R. Cole,[b] and Robert W. Platt[c]

The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients

Daniel Westreich* and Sander Greenland

Revisiting Overadjustment Bias

Haidong Lu
Public Health Modeling Unit and
Department of Epidemiology of
Microbial Diseases
Yale School of Public Health
New Haven, CT

Stephen R. Cole
Department of Epidemiology
Gillings School of Global Public Health
University of North Carolina at Chapel Hill
Chapel Hill, NC

Robert W. Platt
Departments of Pediatrics and Epidemiology
Biostatistics and Occupational Health
McGill University
Montreal, QC, Canada

Enrique F. Schisterman
Epidemiology Branch
Division of Intramural Population Health
Research
Eunice Kennedy Shriver National Institute
of Child Health and Human Development
Bethesda, MD

- This has been provided many groundbreaking insights!

1. A step change in our understanding of selection bias
2. A step change in our understanding of overadjustment bias (and the Table 2 Fallacy)
3. Solved multiple paradoxes!

The Birth Weight "Paradox" Uncovered?

Sonia Hernández-Díaz[1,2], Enrique F. Schisterman[3], and Miguel A. Hernán[1]

Does selection bias explain the obesity paradox among individuals with cardiovascular disease?

Hailey R. Banack MA*, Jay S. Kaufman PhD

The spectre of Berkson's paradox
Collider bias in Covid-19 research

Lord's 'paradox' explained: the 50-year warning on the use of 'change scores' in observational data

*Peter WG Tennant[1,2,3], Georgia D Tomova[1,2,3], Eleanor J Murray[4], Kellyn F Arnold[1], Matthew P Fox[4], Mark S Gilthorpe[3,5]

- These are not trivial achievements!
- Until very recently, overadjustment bias and Table 2 Fallacy were *very* common
- E.g. this scoping review of 421 studies in oral health journals published from 2013-2018 found 45% committed Table 2 Fallacy!

UNSOLICITED SYSTEMATIC REVIEW

Community Dentistry and Oral Epidemiology WILEY

A scoping review of Table 2 fallacy in the oral health literature

Aderonke A. Akinkugbe[1,2,3] | Alyssa M. Simon[2] | Erica R. Brody[4]

[1]Department of Dental Public Health and Policy, School of Dentistry, Virginia Commonwealth University, Richmond, VA, USA

[2]Division of Epidemiology, Department of Family Medicine, School of Medicine, Virginia Commonwealth University, Richmond, VA, USA

[3]Institute for Inclusion, Inquiry and Innovation, Virginia Commonwealth University, Richmond, VA, USA

[4]Health Sciences Library, Richmond, VA, USA

Correspondence
Aderonke A. Akinkugbe, Department of Dental Public Health and Policy, Virginia Commonwealth University. 1101 East Leigh Street, Richmond, VA 23298-0566.
Email: aaakinkugbe@vcu.edu

Funding information
National Institutes of Health/National Institute of Dental and Craniofacial Research, Grant/Award Number: R03DE028403 and L40DE028120

Abstract

Background: Coined by Westreich and Greenland in 2013, Table 2 fallacy refers to the practice of reporting estimates of the primary exposure and adjustment covariates derived from a single model on the same table. This study seeks to describe the extent to which Table 2 fallacy is present in the oral health literature and provide recommendations on presenting findings from multivariable-adjusted models and/or interpretation of adjustment covariate estimates that are not the primary exposure.

Methods: We conducted a scoping review in PubMed and Scopus of human observational studies published in 4 oral health journals (JDR-CTR, CDOE, JPHD, BMC Oral Health) starting in 2013 until the end of 2018. The resulting articles were exported into Excel and were either included or excluded for full-text review based on six criteria. After categorizing the articles, we exported and summarized the results in SAS.

Results: A total of 1358 articles were initially screened of which 937 articles were excluded based on title or abstract for being animal studies, systematic reviews or meta-analysis, prediction models or descriptive studies. The remaining 421 articles were eligible for full text reviewed of which, 189 (45%) committed Table 2 fallacy. The prevalence of table 2 fallacy appears high in the oral health literature.

Conclusions: The problem of presenting multiple effect estimates derived from a single model in the same table is that it inadvertently encourages the reader to interpret all estimates the same way, often as total effects. Implications and recommendations are discussed.

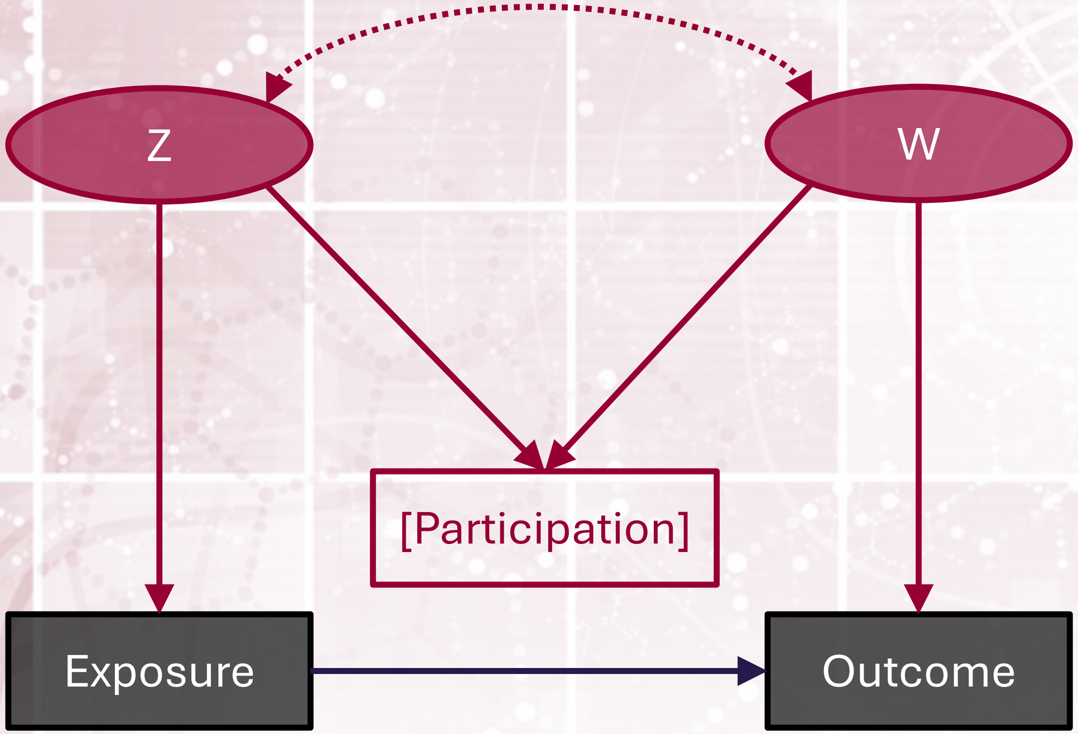
- As recently as 2020, UK Biobank still had the following statement on their website!

biobank uk
Improving the health of future generations

  - "*UK Biobank is not representative of the general population... with evidence of a 'healthy volunteer' selection bias... However, the large sample size and heterogeneity of exposure measures allow for valid...inferences of associations between exposures and health outcomes that are generalizable to the wider population*"

- They even advised users to add the following to their publications :
  - "*Valid assessment of exposure-disease relationships... do not require participants to be representative of the population*"

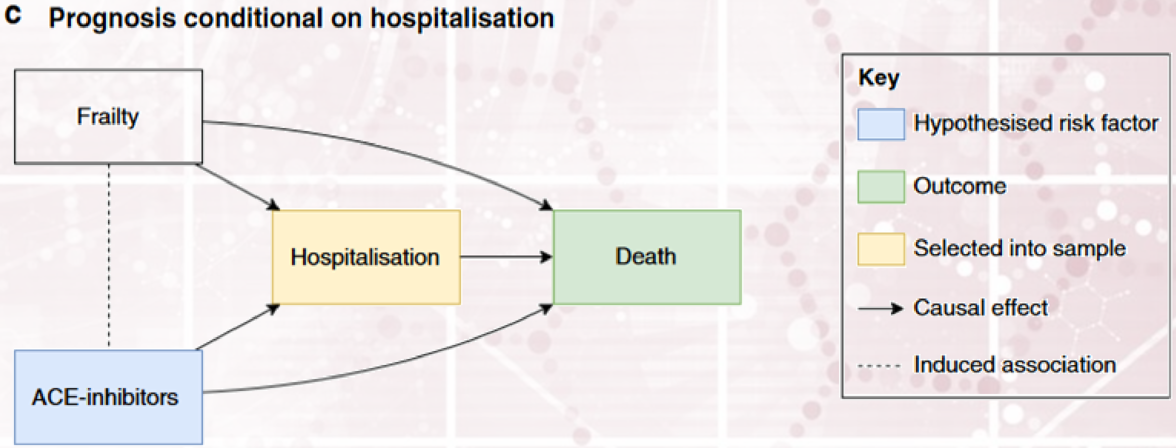
- A simple M-bias DAG explains how non-random selection can distort exposure-outcome relationships in prospective studies



- This knowledge became widely shared during COVID-19 pandemic, when DAGs were used to explain selection bias of risk-factor/severity studies

Collider bias undermines our understanding of COVID-19 disease risk and severity

Gareth J. Griffith[1,2,4], Tim T. Morris[1,2,4], Matthew J. Tudball[1,2,4], Annie Herbert[1,2,4], Giulia Mancano[1,2,4], Lindsey Pike[1,2], Gemma C. Sharp[1,2], Jonathan Sterne[2], Tom M. Palmer[1,2], George Davey Smith[1,2], Kate Tilling[1,2], Luisa Zuccolo[1,2], Neil M. Davies[1,2,3] & Gibran Hemani[1,2,4]



- Many domains have struggled with apparently paradoxical findings
- Thanks to DAGs, these are being resolved at unprecedented pace
- Porta et al 2015: "*The current deconstruction of paradoxes is one among several signs that a profound renewal of methods for clinical and epidemiological research is taking place*"
- These revelations promise real-world impact, by ending confusion and unlocking new models of understanding

Eur J Epidemiol (2015) 30:1079–1087
DOI 10.1007/s10654-015-0068-8

CrossMark

ESSAY

The current deconstruction of paradoxes: one sign of the ongoing methodological "revolution"

Miquel Porta[1,2,3] · **Paolo Vineis**[4,5] · **Francisco Bolúmar**[3,6,7]

Received: 23 March 2015 / Accepted: 4 July 2015 / Published online: 12 July 2015

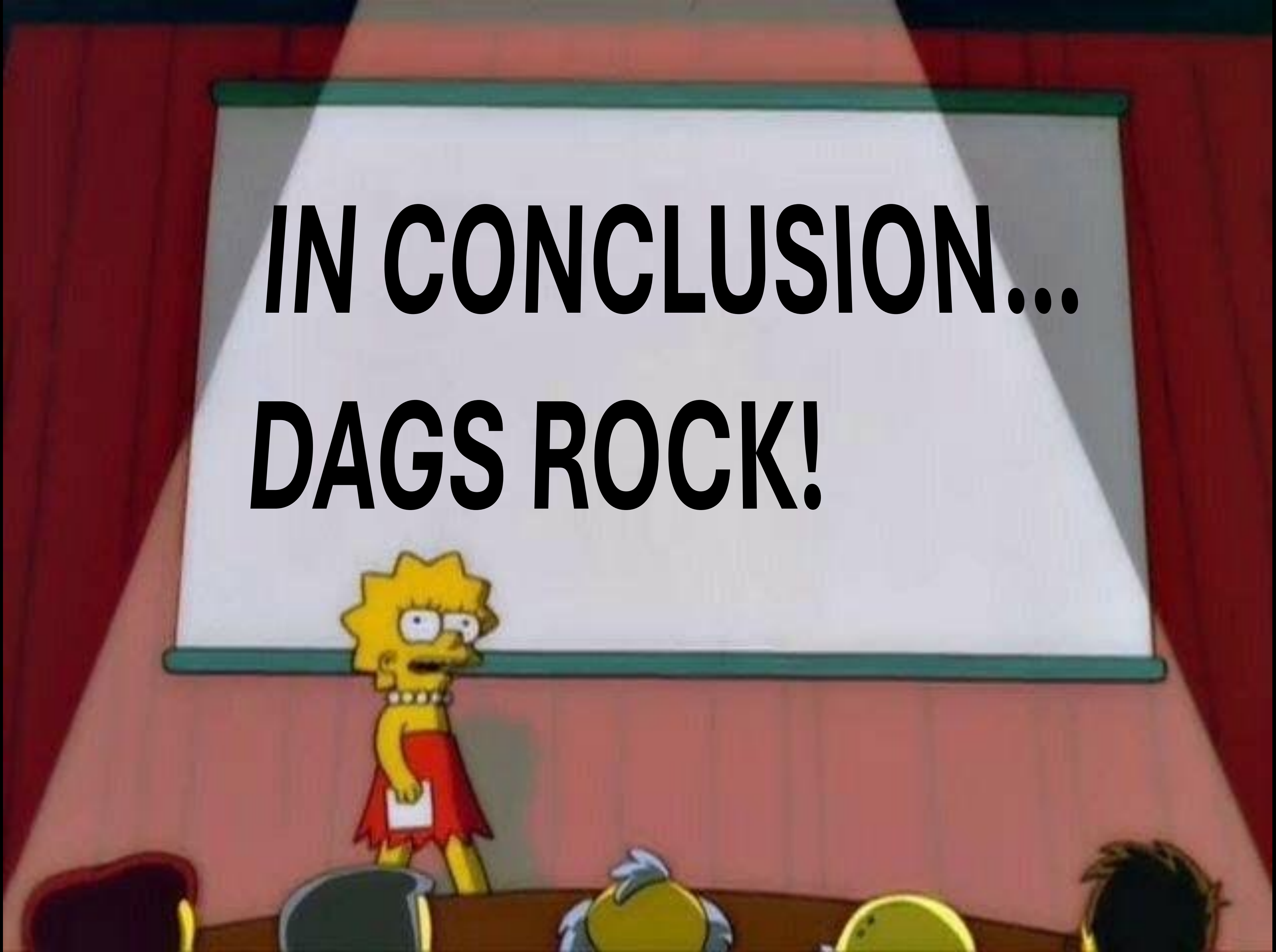
Abstract The current deconstruction of paradoxes is one among several signs that a profound renewal of methods for clinical and epidemiological research is taking place; perhaps for some basic life sciences as well. The new methodological approaches have already deconstructed and explained long puzzling apparent paradoxes, including the (non-existent) benefits of obesity in diabetics, or of smoking in low birth weight. Achievements of the new methods also comprise the elucidation of the *causal structure* of long-disputed and highly complex questions, as Berkson's bias and Simpson's paradox, and clarifying reasons for deep controversies, as those on estrogens and endometrial cancer, or on adverse effects of hormone replacement therapy. These are signs that the new methods can go deeper and beyond the methods in current use. A major example of a highly relevant idea is: when we condition on a common effect of a pair of variables, then a spurious association between such pair is likely. The implications of these ideas are potentially vast. A substantial number of apparent paradoxes may simply be the result of *collider biases*, a source of selection bias that is common not just in epidemiologic research, but in many types of research in the health, life, and social sciences. The new approaches develop a new framework of concepts and methods, as collider, instrumental variables, d-separation, backdoor path and, notably, Directed Acyclic Graphs (DAGs). The current theoretical and methodological renewal—or, perhaps, "revolution"—may be changing deeply how clinical and epidemiological research is conceived and performed, how we assess the validity and relevance of findings, and how causal inferences are made. Clinical and basic researchers, among others, should get acquainted with DAGs and related concepts.

DAGs promised to:

- **Help identify variables that require measuring and adjusting**
  - *Thousands of researchers are using DAGs to help identify variables for adjustment*
  - *They make it much easier to seek diverse input into study design*

- **Help with communicating results and assumptions**
  - *They are useful for explaining modelling decisions and highlighting sources of bias*
  - *They make implausible assumptions much clearer!*

- **Help with understand and avoiding different types of bias**
  - *They have revolutionised our understanding of selection and overadjustment biases*
  - *They are resolving confusions and paradoxes in countless domains*


IN CONCLUSION...
DAGS ROCK!

The Alan Turing Institute

LEEDS Institute for Data Analytics

UNIVERSITY OF LEEDS

Have DAGs fulfilled their promise?

Dr Peter WG Tennant

University of Leeds, Leeds UK

Alan Turing Institute, London, UK

@PWGTennant

WCE

WORLD CONGRESS OF EPIDEMIOLOGY 2024

- According to Greenland, Pearl, and Robins (1999) and Shrier and Platt 2008, DAGs had three main promises:

- **1) Helping identify variables that require measuring and adjusting to estimate a causal effect**

- **2) Communicating results**

- **3) Helping to understand and avoid different types of bias**

PROMISE 1: HELPING IDENTIFY VARIABLES

- Thousands of scientists are using DAGs to ‘help identify variables’ for adjustment in studies and protocols

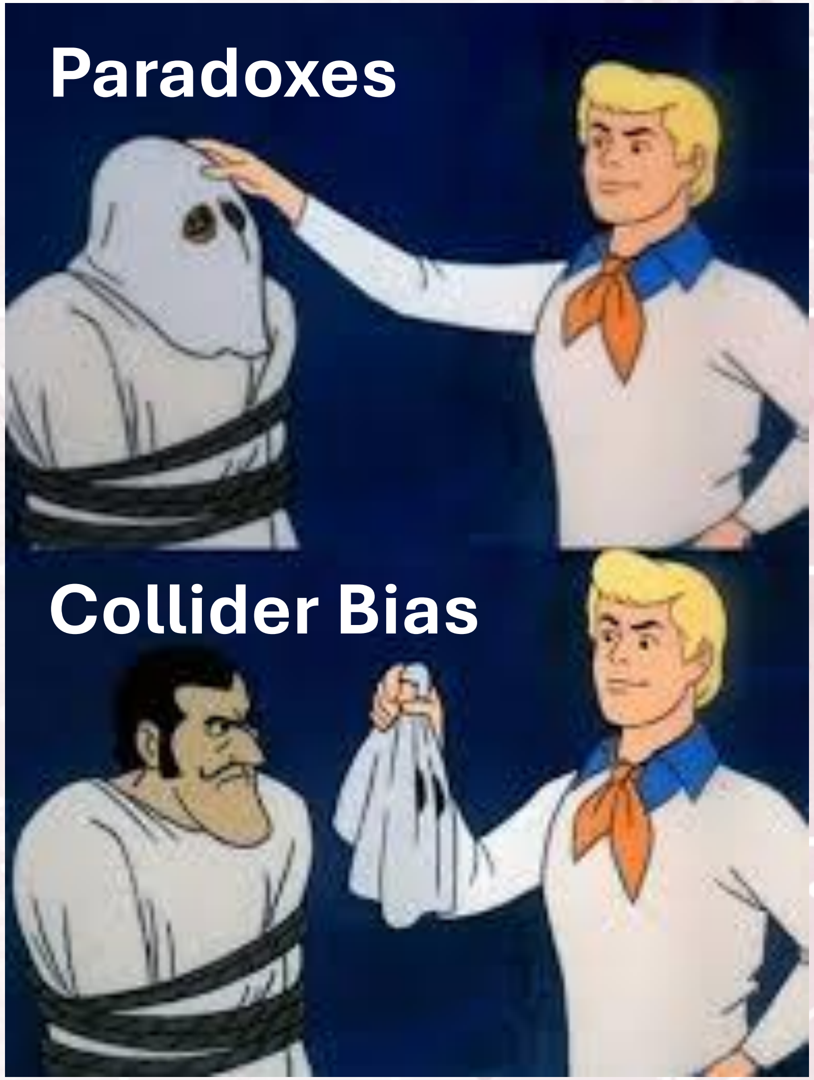
- **This brings two substantial benefits:**

1. It makes the process of identifying and selecting variables for adjustment much more transparent

2. It supports wider stakeholder input

- Using DAGs offer a huge improvement in transparency
- Helps to highlight residual sources of bias
- Helps explain the context and key modelling decisions
- Makes (implausible) assumptions much more visible!

- DAGs have revolutionised our understanding of various forms of error and bias
- This is revolutionising our discipline, solving various paradoxes and revealing new insights
- We understand selection bias and overadjustment bias better than ever before
- Many tricky issues and concepts can be explained to large audiences using DAGs




VOTE YES!