Essay

Using Directed Acyclic Graphs (Dags) to Represent the Data Generating Mechanisms of Disease and Healthcare Pathways: A Guide for Educators, Students, Practitioners and Researchers

George Ellison *

Posted Date: 8 October 2022

doi: 10.20944/preprints202210.0085.v1

Keywords: Directed Acyclic Graph; DAG; confounding; collider bias; epistemology; inferential statistics

Using Directed Acyclic Graphs (DAGs) to Represent the Data Generating Mechanisms of Disease and Healthcare Pathways: A Guide for Educators, Students, Practitioners and Researchers

George TH Ellison

Key words: Directed Acyclic Graph; DAG; confounding; collider bias; causal inference; epistemology

1 Introduction

1.1 Bridging the Divide Between Clinical and Statistical Expertise Using Directed Acyclic Graphs

The transition from 'knowledge-consumer' to 'knowledge-producer' can require considerable cognitive agility. It can be all the more challenging when the methodological techniques necessary to *produce* knowledge rely on: the prior *consumption* of substantial amounts of very different, and very specialised/technical, types of knowledge; and the confidence of those involved that they have the 'cognitive competence' required to assimilate and apply this [1]. Bridging the divide between disparate forms and levels of expertise is necessary to ensure that advances in any one area can support improvements in understanding and application by *non*-specialists elsewhere; and by generalists and specialists in *several* (or *other*) areas of expertise. Challenges in the building of such bridges exist in many different contexts — not least in the public understanding of advanced scientific techniques (and the 'informed consumption' of the findings, insights and outputs these techniques generate; e.g. [2, 3]). Moreover, many of these challenges are intensifying as a result of the increasing levels of specialism required to attain the levels of knowledge, understanding and technical competence required to break new ground in any given discipline, or on any given topic [4]. For these reasons, the duration and intensity of professional education programmes can struggle to keep up with the rapid pace of advances in knowledge and technology. Nowhere is this more keenly felt than in the fields of statistics, applied mathematics and data science, where a recent step change in the volume of deliberately, routinely and incidentally digitised data, and in the computational capacity required to

GTH Ellison
Centre for Data Innovation, Faculty of Science & Technology, University of Central Lancashire, Preston, UK; Centre for Academic Technologies, University of Johannesburg, Johannesburg, Gauteng, Republic of South Africa. E-mail: gthellison@uclan.ac.uk; ORCID: 0000-0001-8914-6812

Pages 61-101; Chapter 6 in: Medeiros Mirra RJ and Farnell D (ed.s) Teaching Biostatistics in Medicine and Allied Health Sciences. Springer Verlag: 2023; ISBN: 978-3-031-26009-4). DOI: 10.1007/978-3-031-26010-0_6.

collate, manage and analyse these data, has led to a proliferation of advanced analytical techniques, and a deluge of findings, insights and claims — each of which require expert evaluation and thoughtful consideration prior to their interpretation and application.

For specialists involved in producing knowledge for public consumption — or for use by specialists in *other* domains and by generalists working across *multiple* domains — bridging the divide between the techniques they use, the evidence these generate, and the contribution they might make to strengthening and extending the decisions and practice of others, has led to growing interest in the science of 'exposition', 'translation' and 'implementation'. This has involved the development and evaluation of innovative approaches, techniques and devices designed to facilitate the rapid dissemination, acquisition and integration of novel perspectives; and thereby extend the impact and insight that such perspectives might bring to bear *elsewhere* [4, 5]. These approaches, techniques and devices include: improving the accessibility of specialist or technical language (e.g. [6]; this volume); providing accessible or 'plain language' summaries and 'visual abstracts' of 'key findings' [7]; creating so-called 'user guides' to explicate the 'experiential context' of specialist applications (see for example [8]; and https://mathusersguides.com/); and using a variety of visualisation practices to simplify and demystify the techniques involved and the information these generate [2, 9–11].

Unfortunately, what we know from the use of terms, idioms, similes, analogies and metaphors that cross over from popular (or generalist) to professional (or specialist) 'experiential contexts' is that these risk conflating their very different, context- specific meanings (and very different semantic etymologies) in ways that can mislead, confuse and obfuscate (rather than elucidate, clarify or enlighten). Examples include terms such as: 'significance' and 'prediction', which mean very different things within inferential statistics and popular discourse [12, 13]; and 'machine learning' and 'artificial intelligence', which were ostensibly coined to provide an insight into the complex analytical procedures involved, but have encouraged the widespread popular misconception that both are autonomous, uncontrolled and potentially uncontrollable [14].

Nonetheless, translation appears more successful when the terminology and techniques used are derived from (and remain pertinent, relevant and salient to) the specialist 'producers' involved (as is the case with terms such as 'coefficient' and 'effect size' [15] — although the term 'effect' has itself become a problematised concept within evidence-informed decision-making). Indeed, many of the most successful strategies adopted to-date seem to be those that co-opt a specialist term or device and repurpose this as a medium through which the principles and procedures involved can be more faithfully translated for a wider audience (including the lay public, non-specialists, generalists and specialists in unrelated fields).

This has been the rationale behind the use of directed acyclic graphs (DAGs) to facilitate the acquisition of secondary statistical skills by third year medical students [16–18], since these diagrams help bridge the divide between: the aetiological knowledge and professional/operational experience on which clinical training and

practice depend; and the more abstract epistemological and analytical considerations required to extract robust statistical insight from health and healthcare data. DAGs lend themselves to ancillary statistical training precisely because they provide *nonparametric* representations of both: disease and healthcare pathways; and the underlying 'data generating mechanisms' involved — i.e. the mechanisms on which any *parametric* insights generated using inferential statistics depend. As such, DAGs should make inferential statistics much more accessible for those healthcare students, practitioners or researchers who lack high school training or formal qualifications in statistics; and particularly those with little aptitude for, or interest in, the mathematical concepts involved (including those who suffer from 'numerophobia' or 'mathematical anxiety' [19–21]).

1.2 Directed Acyclic Graphs as Epistemological and Educational Tools for Inferential Statistics

Healthcare students, practitioners and researchers unacquainted with DAGs are likely to find them comfortingly familiar and remarkably accessible, since they closely resemble the path diagrams and flowcharts that have become popular for summarising a huge variety of sequential and contingent processes, including disease and healthcare pathways (see, for example: https://www.wikipathways.org/ [22]; and https://cks.nice.org.uk/ [23]; and [24]; see also Fig. 1, below). Indeed, although DAGs are *directed* rather than *directive* — and do not include (or require) the symbols that formal flowcharts use as guides for navigating between 'start' and 'end' points [26] — their use of unidirectional arrows (or 'arcs') to represent causal paths between successive variables (or 'nodes') mirrors the directional 'connectors' that flowcharts use to move from 'start' to 'finish' via a sequence of 'inputs', 'processes' and 'outputs' contingent on 'decisions' made at key points along the way.

Nonetheless, a more critical distinction between DAGs and flowcharts is the former's 'acyclicity'. This requires that no variable can act as its own cause — either *directly* (through a single causal path that starts from, and ends with, the variable concerned) or *indirectly* (through a causal pathway involving a sequence of causal paths between successive variables that eventually loop back to the original variable). Instead, DAGs are inexorably 'progressive' — as befits their temporal disposition — and reflect pathways that link successive pairs of variables over time through a sequence of causal paths between each of the variables acting as a preceding probabilistic cause, and each of the subsequent variables acting as one of their subsequent consequences. It is therefore on the basis of these two key principles — *unidirectionality* and *acyclicity* — that *directed* (and) *acyclic* graphs derive their name; and why DAGs constitute 'principled' diagrammatic representations informed by these two key rules (or 'principles'). More formally, DAGs constitute a subset of causal path diagrams that offer theoretical representations of the known, uncertain or entirely speculative causal relationships thought to be operating within any given context; and thereby offer clear, detailed

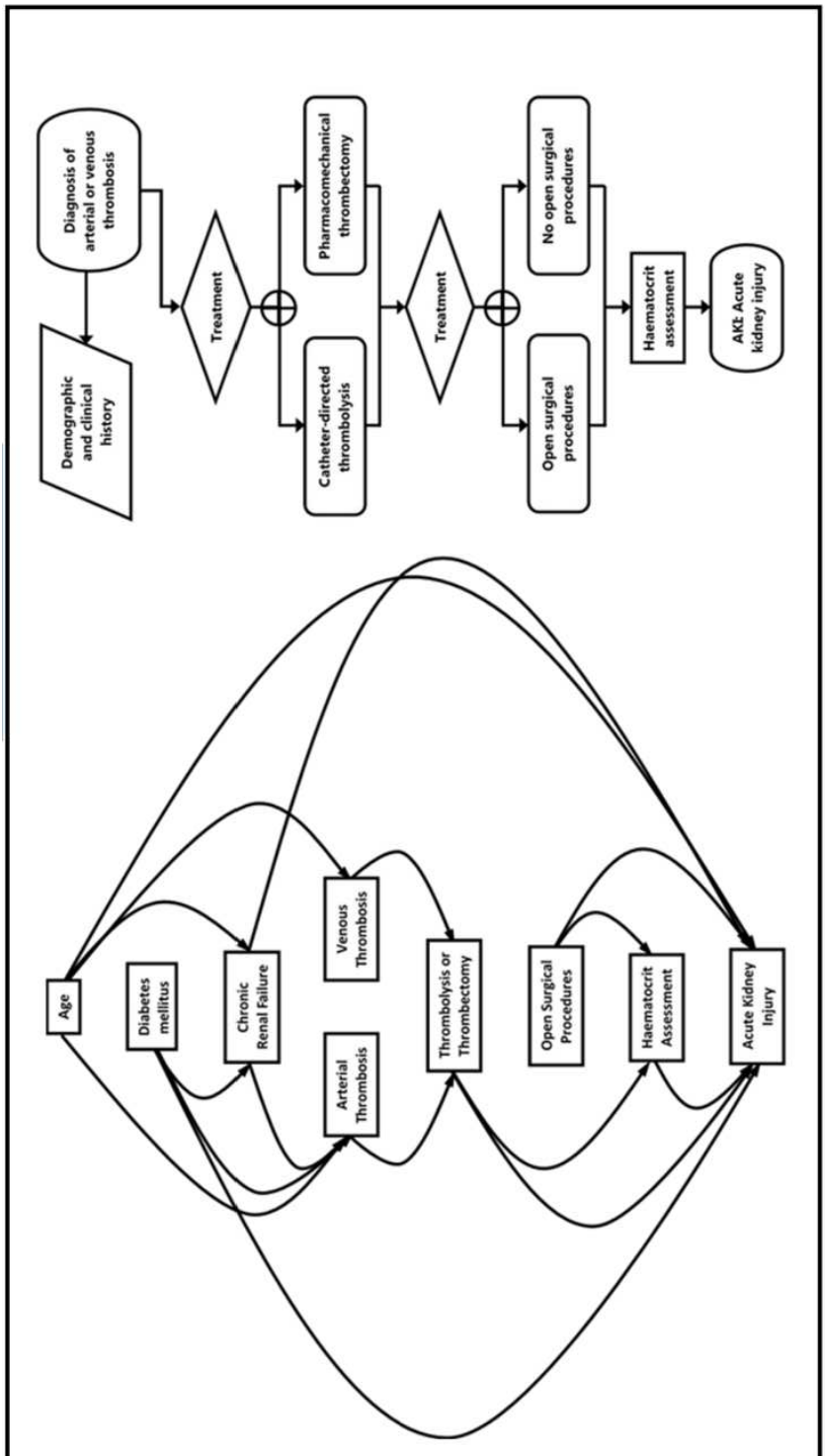


Fig.1 The similarities and differences between a causal path diagram (below, left), redrawn past-to-present/top-to-bottom from the DAG provided by Escobar et al. ([25 ; Figure 7: 244 PA; 29 AAM), to examine the causal effect of two alternative clinical interventions on acute kidney injury in hospitalised thrombosis patients; and a formal flowchart (below, right) drawn to describe the sequence of clinical decisions/outcomes involved in the healthcare these patients received

and principled descriptions of the contributions that theoretical understanding, empirical evidence, or more tentative (or speculative) hypotheses can make to epistemological models of the data generating mechanisms involved [27].

While DAGs, like flowcharts, only provide *nonparametric* (that is qualitative rather than quantitative) representations of the pathways involved in the processes they describe, both DAGs and flowcharts have *parametric* (i.e. quantitative) implications for the influence that preceding phenomena can exert on subsequent events (be these 'inputs', 'outputs', 'processes' and 'decisions' — as in a flowchart; or, more simply, the variables acting as successive causes and consequences — as in a DAG). This is because the arrangement of causal paths (and in particular, the *omission* of temporally plausible causal paths) imposes strict parametric constraints on the distributional properties of the data sets generated by the mechanistic pathways they represent.

DAGs can therefore be drawn without the need to consider whether the causal pathways involved are strong or weak; positive or negative; precise or diffuse. They can also be drawn with little (if any) knowledge of the: parametric characteristics of the variables concerned; or the distributional properties these pathways impose thereon. Yet DAGs successfully map the parametric *consequences* of the data generating mechanisms they describe, and thereby facilitate the *statistical* analysis of the datasets these mechanisms generate — an invaluable benefit for inexperienced or non-specialist analysts keen to derive robust statistical inference from the underlying causal relationships involved. This is what makes DAGs such powerful translational tools for bridging the divide between aetiological knowledge, professional/operational experience, and the skills required for inferential statistics. Put simply, DAGs can be drawn with little concern for statistical or mathematical expertise but, once drawn, make it much easier to conduct statistical analyses, and acknowledge, mitigate and (potentially) eliminate potential risks of analytical or inferential bias involved therein.

1.3 The Vulnerability of Directed Acyclic Graphs to Imperfect Aetiological Knowledge and Professional/ Operational Experience

Despite the translational benefits of DAGs — which, as we have seen, stem from their dual utility as models of disease and healthcare pathways, and of the underlying data generating mechanisms these entail — drawing DAGs that *accurately* represent the processes involved can be challenging. While one might expect this to be least troublesome for healthcare pathways — since these are ostensibly the artefactual products of (known) operational designs and (deliberate) professional practices — such pathways often depend upon a complex interplay of contextual (structural, social and cultural) factors that can be uncertain, obscure and very different to what was *intended*, and therefore extremely challenging to accurately (or even adequately) describe. Likewise, despite ongoing advances in theoretical understanding, aetiological knowledge and technical expertise, few (if any) disease

pathways are *comprehensively* understood. And even where such understanding, knowledge and expertise exist, the processes involved are often contingent on a similarly complex interplay of contextual factors that influence their incidence, presentation, observation, recognition, diagnosis, treatment and prognosis. Such factors can therefore make *both* healthcare *and* disease pathways susceptible to uncertainty and misunderstanding, even by those with advanced training, first-hand experience or specialist expertise. Indeed, the level of uncertainty involved is likely to be higher still amongst healthcare students and less experienced practitioners and researchers whose understanding will depend almost entirely on the extent of the knowledge they have received in their formal training (together with whatever, potentially limited and potentially partial, exposure they have had to the disease and healthcare pathways concerned).

Moreover, even when substantial relevant knowledge, insight and experience of the variables and causal pathways *is* available, any DAGs based thereon will remain susceptible to 'information bias' [28, 29] and 'anchoring bias' [30, 31]. This is because their content and structure will be: constrained by what is likely to be incomplete or inaccurate *information*; or *anchored* to limited prior knowledge or imperfect biomedical/organisational paradigms. DAGs will also be vulnerable to the so-called 'availability heuristic' [32, 33] whenever they (sub)consciously preference information that is most amenable to observation or measurement (such as including only those variables for which data are already, routinely or readily *available*). Under these circumstances, any such DAGs will offer limited scope for statistical insight free from analytical or inferential bias whenever they ignore or overlook the important role that limited information (such as omitted or unmeasured variables) might play therein.

1.4 The Utility of Temporal Logic When Teaching, Learning and Using Directed Acyclic Graphs

Fortunately, the temporal-causal nature of disease and healthcare pathways mean that the DAGs used to represent these (and the underlying data generating mechanisms involved) can often be drawn using temporal logic alone. In fact, provided the temporal sequence of all relevant variables (or, rather, the phenomena, characteristics and features these variables represent) is known or can be inferred with a fair degree of certainty, temporality alone will dictate which causal paths are *impossible* and which might at least be *plausible*. This level of certainty can be achieved simply on the basis that no subsequent variable can *possibly* cause any preceding variable, while any preceding variable can be considered a *plausible* (or so-called 'probabilistic') cause of any subsequent variable, albeit in the absence of *definitive* evidence to the contrary (such as the deliberate, random allocation of preceding characteristics amongst participants who do and do not receive an intervention — as is the case within a randomised controlled trial).

For these reasons, once one has specified a causal pathway of interest (what Aneshensel [34] helpfully called the 'focal relationship') — comprising the

plausible, probabilistic causal path between a preceding cause (or 'exposure' variable) and a subsequent consequence (or 'outcome variable') — all that is required to draw a *preliminary* DAG using temporal logic is to:

- **Step 1** — compile a comprehensive list of any phenomena, characteristics or features (and their associated variables) considered relevant to the focal relationship;
- **Step 2** — arrange all of the variables identified at Step 1 in the most plausible temporal sequence on the basis of when the phenomena, characteristics and features involved (as and when each of these were measured) were most likely to have occurred; and
- **Step 3** — include *all* of the plausible, probabilistic causal paths from any preceding variable to all subsequent variables (i.e. as assessed at Step 2, above), only omitting those paths where there is *definitive* evidence these do not or cannot exist.

The aim of this chapter is to illustrate how educators can apply these three (and seven further, simple) steps to strengthen the analytical competence and statistical confidence of healthcare students, practitioners and researchers using DAGs informed by temporal logical. The sections that follow include a worked example involving the critical appraisal of a published clinical study [25] to demonstrate how each of the tasks required to generate a temporally robust DAG can also be used to: evaluate the analytical decisions made during applied healthcare research; and inform the decisions required when selecting the datasets and statistical adjustments required to eliminate, mitigate or (at the very least) acknowledge the risk of bias in similar, real-world, non-randomised studies (i.e. those dependent on routinely or deliberatively collected observational data).

Escobar et al.'s [25] study was chosen as the basis for this chapter's worked example to represent the 144 studies reviewed by Tennant et al. [35], each of which had used (and shared) DAGs to inform the design of their applied health research. Like many of these studies, Escobar et al. [25] drew on a combination of routinely and prospectively collected sociodemographic and clinical data from a discrete subset of patients (in this instance, following a diagnosis for thrombosis), to estimate the strength, direction and precision of a clearly specified focal relationship (in this instance, between the type of treatment received/allocated and a marker of subsequent kidney damage). In common with many of the other studies reviewed, Escobar et al.'s [25] analyses involved multivariable statistical modelling in which the adjustment sets used were informed by a DAG that had been drawn to represent the aetiological and healthcare pathways involved, and determine which variables might act as potential confounders (and should therefore be included in the multivariable adjustment sets used). However, in contrast to some of the other articles that Tennant et al. [35] reviewed, Escobar et al. [25]: contained all of the information required to evaluate the key analytical decisions made; was considered accessible to non- specialists; and included just one clearly drawn (and therefore unambiguous) DAG, containing a manageable number of discrete variables ([25];

Figure 7: page 244 of the published article [PA]; page 29 of the Author Accepted Manuscript [AAM]).

The first three of these steps were used to generate the DAG presented in Fig. 2 (below).

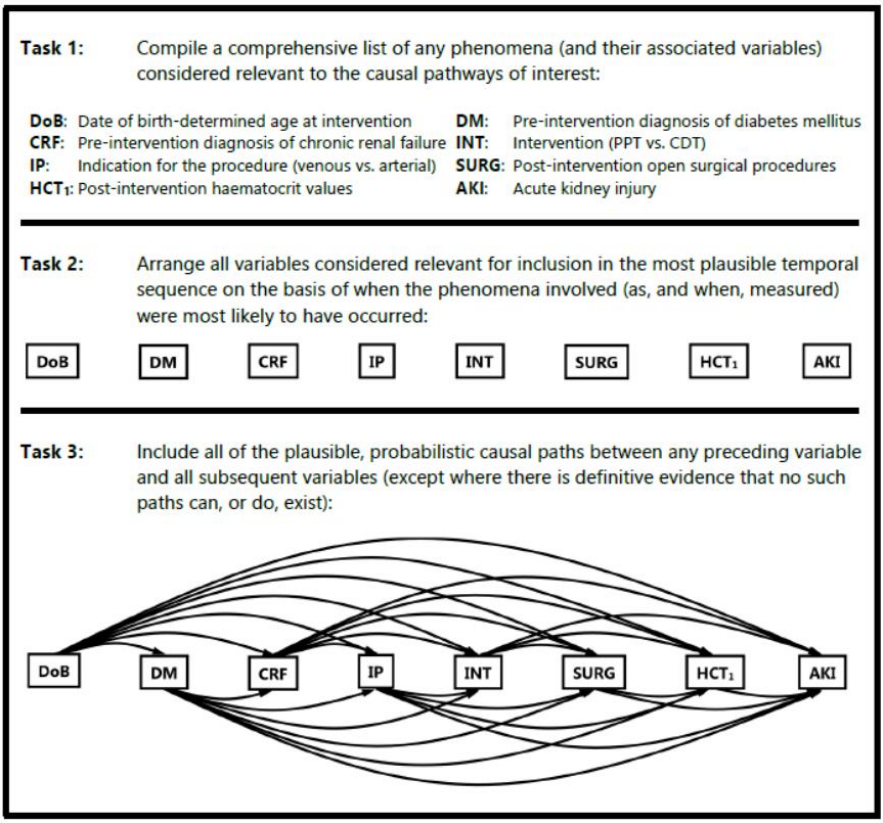


Fig. 2 An illustration of the first three tasks involved in drawing preliminary DAGs based on temporal logic, using the eight discrete variables included in the DAG provided by Escobar et al. ([25]; Figure 7: 244 PA; 29 AAM), to examine the causal effect of percutaneous pharmacomechanical thrombectomy (PPT) vs. catheter-directed thrombolysis (CDT) on acute kidney injury in thrombosis patients. (See: DAGitty)

This comprises a reformatted and redrawn version of the DAG provided by Escobar et al. ([25]: Figure 7: 244 PA; 29 AAM) which was: generated after identifying the eight discrete variables from the nine nodes the authors had included in their original DAG; and based around the focal relationship between INT: the clinical procedure/intervention allocated/received (the specified exposure variable), and AKI: acute kidney injury (the specified outcome variable). At this stage, it is assumed that this list of eight discrete variables comprised those the authors considered both important and sufficient (if not necessarily *comprehensive*) to describe the relevant causal pathways of interest, and the data generating mechanism involved. In this regard, it is worth pointing out that perhaps as many as 75 different

phenomena, characteristics, features, and their associated variables were mentioned or cited in the three main sections (Introduction, Methods, Results and Discussion), two Tables (I and II) and seven Figures (1–7) of Escobar et al.'s [25] article. These have been summarised and tabulated in the appendices to this chapter (see Appendix 1), and indicate the far larger number of variables the authors themselves considered relevant to their study's rationale, design, context, sample, analyses and interpretation, yet failed to include (either as measured or unmeasured/latent variables) in the DAG they presented in Figure 7: 244 PA; 29 AAM).

It should be clear from Fig. 2 that these first three, preliminary steps should make drawing (and teaching) DAGs both simple and accessible, even for those who have little if any aetiological/professional knowledge, or first-hand operational experience of the disease and healthcare pathways concerned [16]. Moreover, given the risks of anchoring and information bias, it may not be unreasonable to speculate that drawing such DAGs based on temporal logic *alone* might actually generate *more* objective, and perhaps *more* accurate, representations of the pathways concerned than those drawn on the basis of existing biomedical paradigms or professional/ operational understanding (given these are inevitably partial, imperfect or incomplete). Indeed, even when there is a substantial degree of *un*certainty regarding the precise temporal sequence of the variables involved, DAGs informed by temporal logic should still have substantial utility for: prompting epistemological reflection and thoughtfulness; and assessing whether the causal paths considered likely or necessary (on the basis of prevailing knowledge, insight and experience) are consistent with those considered possible (and therefore plausible) on the basis of temporality — something that is remarkably easy to overlook when one is *anchored* to established biomedical/operational paradigms or analytical designs.

2 Drawing Directed Acyclic Graphs Using Temporal Logic

2.1 Challenges Facing the Development of Sufficiently Comprehensive and Temporally Robust DAGs

Nonetheless, while none of the three *preliminary* steps required to generate a DAG on the basis of temporal logic appear particularly taxing, the first two face a number of substantive challenges. This is because compiling a *comprehensive* list of all variables likely to be relevant to any given causal pathway/focal relationship (Step 1), and arranging these in their correct temporal sequence (Step 2), requires very careful consideration of any variables that might be:

- 'omitted' — not least those that are known, measured yet overlooked, but also those that are: known but unmeasured; unknown and therefore unmeasured; and potentially 'unknowable' and therefore unmeasurable and likely to be unacknowledged [36];

- 'time-variant' — and therefore highly dependent on precisely when and how they were measured, and the relevance of the values these measurements provide to the causal effects of their *preceding* values;
- 'non-asynchronous' — such that they are impossible to confidently position before or after one or more other variables, which therefore appear to have occurred at exactly (or essentially) the same point in time; and/or
- 'temporally obscure' — when there is intractable uncertainty regarding their temporal position with regard to one or more of the other relevant variables (and, in particular, in relation to the specified exposure and outcome variables of the focal relationship of interest).

Such variables can pose substantive challenges when drawing a DAG since: the first (omitted variables) include those that are easily overlooked and, as a result, are commonly under-represented within DAGs (see also Appendix 1); the second (time-variant variables) are difficult to operationalise as discrete, temporally positioned phenomena (not least when these actually relate to characteristics or features rather than 'events' per se); the third (non-asynchronous variables) introduce considerable uncertainty regarding their causal relationships with other simultaneously occurring/nonasynchronous variables; and the fourth (temporally obscure variables) can undermine the whole premise of drawing DAGs using, or informed by, temporal logic [16].

The four sections that follow deal with each of these challenges in turn; offering guidance on how they might be addressed through a series of additional steps which move beyond simplistic, *preliminary* DAGs (such as that summarised in Fig. 2, above), to more comprehensive DAGs that acknowledge (and can therefore help mitigate, or even eliminate), the potential impact of omitted, time-variant, non-asynchronous and temporally obscure variables on the accuracy and analytical utility of the data generating mechanisms such DAGs aim to represent.

2.2 Challenges Posed by the Omission of Measured, Unmeasured and Unacknowledged Variables

Compiling a comprehensive (yet finite and manageable) list of variables considered relevant to any given causal pathway(s) will always require some pragmatic decisions — particularly regarding which of the measured (and therefore available) variables, and which of the unmeasured (and unmeasur*able*) variables, might be both *necessary* and *sufficient* to represent the data generating mechanism(s) involved (see also Appendix 1). This is evident from Tennant et al.'s [35] recent review of 144 published DAGs, in which the number of variables included ranged from 3 to 28 with an average of just 12. This may partly reflect the decline in readability that accompanies the inclusion of more than a dozen variables (and their associated causal paths) within a static two-dimensional DAG [16]. However, it also seems likely that (for the most part) the specific variables selected for inclusion in many of these DAGs were either: those for which measurements were readily or

already available; or those considered critical to include on the basis of prior aetiological knowledge (both theoretical and empirical) or professional/operational experience. If so, then the very basis upon which many contemporary DAGs are drawn will be vulnerable to the 'availability heuristic' and 'anchoring' bias, respectively.

Notwithstanding the need to recognise, reduce and (where possible) avoid the risks that such cognitive biases entail [37], it remains common practice to ignore unmeasured variables when drawing DAGs. Indeed, Tennant et al. [35] reported that only a third of the 144 DAGs they reviewed contained one (or more) unmeasured variable(s), while none appeared to have included *unacknowledged* variables — i.e. the 'unknown unknowns' and potentially '*unknowable* unknowns' that fall outside existing knowledge, theory or perception, and are therefore unmeasurable [36]. A failure to consider the potential relevance of such variables when drawing a DAG both replicates and reinforces the narrow gaze imposed by excessive reliance on: (partial) aetiological knowledge; (limited) professional/operational experience; (imperfect and incomplete) biomedical and organisational paradigms; and variables for which measurements are more readily (or are already) available. When no unmeasured variables and no unacknowledged variables are included, the DAGs drawn are much more likely to misrepresent the complexity of the causal pathways involved. They also overlook: the important role(s) that any of the omitted variables (whether known, unknown or unknowable) might play therein; and the potential utility of conceptualising and operationalising such variables within the theoretical and empirical deliberations that DAGs are intended to promote.

For these reasons, all such variables — including those that are measured, unmeasured *and* unacknowledged — should be considered relevant to include when drawing a DAG since all of these can: prompt reflection and generate insight regarding alternative aetiological and operational pathways; and help highlight potential sources of analytical and inferential bias that might otherwise be overlooked. To this end, any 'measured-yet-omitted' and all 'known-but-unmeasured' variables should be positioned wherever they are likely to have occurred within the temporal sequence of variables already selected for inclusion; while hypothetical (multivariable) 'sets' of unacknowledged variables should be distributed throughout the temporal sequence of (measured and unmeasured) variables, and particularly at those points in time where they would pose specific and credible risks of analytical and inferential bias to any subsequent statistical analyses (as described in more detail in Sect. 3.2, below).

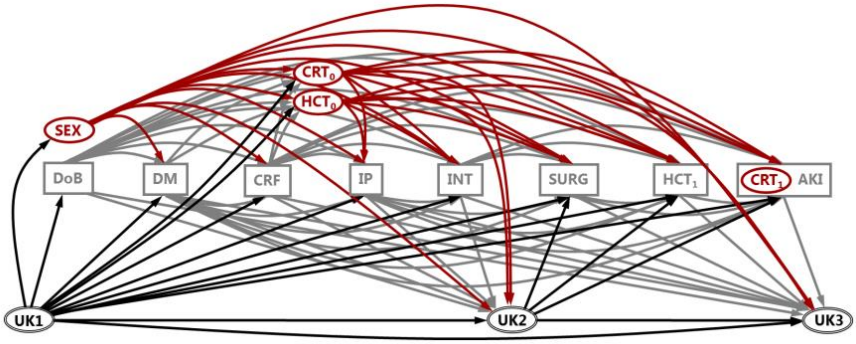


Fig. 3 A further iteration of the DAG presented in Fig. 2 to illustrate how: (1) additional, measured-yet-omitted and known-but-unmeasured/unreported variables (SEX, CRT_0, HCT_0 and CRT_1 — in **red font**); and (2) unknown and hitherto unacknowledged sets of variables (UnKnown 1-3: UK1, UK2, and UK3 — in **black font**) might be added to prompt reflection and generate insight regarding: alternative aetiological/operational pathways; or potential sources of analytical/inferential bias that might otherwise be overlooked; and (3) how variables considered non-asynchronous (such as: DoB and SEX; and CRT_0 and HCT_0) might best be located (as events occurring simultaneously with no causal pathways operating between them). (See: DAGitty)

To illustrate how unmeasured and unacknowledged variables might be usefully incorporated into a DAG, Fig. 3 extends the worked example (Fig. 2 — based on the DAG provided by Escobar et al. [25]: Figure 7: 244 PA; 29 AAM) to reflect two further steps:

- **Step 4** — include all relevant variables (including any measured-yet-omitted, and all known-but-unmeasured, variables); and
- **Step 5** — include hypothetical sets of potentially unacknowledged (and therefore unknown, unmeasured and potentially unmeasurable) variables at key points within the DAG to prompt consideration and reflection regarding the risks of analytical and inferential bias these might pose.

The first of these steps (Step 4) identified three additional measured variables (the sex of the patients examined — SEX; their baseline creatinine levels — CRT_0; and their baseline haematocrit values — HCT_0), all of which had been reported in Table 1 of Escobar et al. (25), yet all of which were omitted from the published DAG and the multivariable analyses the DAG was designed to inform (para. 1: 241 PA; 21 AAM). It also identified a fourth measured, but unreported (and therefore omitted), variable, namely: the follow-up creatinine values of the patients examined (CRT_1; recorded within 72 hours of INT) which, though omitted from the DAG and multivariable analyses, nonetheless informed (and were integral to) the diagnosis of AKI.

Meanwhile, the second of these steps (Step 5) involved: positioning three hypothetical sets of variables (UnKnown 1-3: UK1; UK2; and UK3): prior to the very first of the measured variables (the date of birth-determined age of the thrombosis patients; DoB); after the very last of the measured variables (acute

kidney injury: AKI); and in-between the specified exposure (INT) and specified outcome (AKI; i.e. along the causal path of the focal relationship). As discussed in Sects. 2.5 and 3.2, below, these are the locations within a DAG where other, 'third party' variables (so-called 'covariates') might pose very particular risks of analytical and inferential bias when estimating the strength, direction and precision of any given focal relationship.

The inclusion of these additional (measured yet omitted) variables, and the three sets of (hypothetical, unmeasured and potentially unmeasurable) variables within Fig. 3 — together with their associated probabilistic causal paths — demonstrates how much more complex DAGs can become when compared to those drawn solely using a small number of measured, available, accessible and/or salient variables (as in Fig. 2, and the original DAG provided by Escobar et al. [25]: Figure 7: 244 PA; 29 AAM). This added complexity nonetheless has substantial utility for: tempering excessive confidence that any such models might comprehensively represent the multitude of aetiological and operational pathway(s) involved, and the data generating mechanism(s) these reflect; as well as revealing and exploring alternative target variables and causal pathways that might warrant further investigation and statistical modelling as possible targets for subsequent intervention. At the same time, the inclusion of additional unmeasured and unacknowledged variables can also be invaluable for considering, evaluating and identifying potential sources of analytical or inferential bias, so that these are not ignored or overlooked, and can instead be eliminated, mitigated or (at the very least) acknowledged (see Sect. 3.2, below).

2.3 Challenges Facing the Temporal Positioning of Time-Variant Variables

Notwithstanding the potential benefits of including unmeasured variables within any given DAG, their inclusion is nonetheless predicated upon the ease and accuracy with which the temporal sequence of these (and other, measured) variables can be determined. Likewise, adding hypothetical sets of unacknowledged variables at those points in the temporal sequence where they might best prompt consideration of, and reflection on, the substantive risk of analytical and inferential bias that such variables might then pose is also predicated upon establishing a temporal sequence of measured and unmeasured variables amongst which these additional sets of variables might then be positioned. In practice, determining (or inferring) the temporal sequence of any group of variables will largely depend upon whether the variables concerned represent phenomena, features or characteristics that are 'time-invariant' or 'time-variant' — that is, whether these variables refer to discrete, one-off events, happenings or occurrences (hence 'time-*in*variant'); or are susceptible to change over time (hence 'time-*variant*').

It should be relatively straightforward to identify when a time-invariant variable occurs (or occurred) since this can be traced to the associated event concerned. However, wherever there is substantive variation in the precise point in time at which such events occur these variables can also present as essentially time- *variant*

at the population level, making it necessary to estimate the most likely or commonest point in time *at* which (or *by* which) the phenomena concerned will have occurred for *most* (or *all*) of the cases examined (be these study participants or other entities). In a similar fashion, identifying when a time-variant variable occurs (or occurred) is likely to prove much more challenging, because although their values are dependent upon *when* and *how* they were measured, the relevance of these values to the causal mechanisms involved may actually depend upon values attained much *earlier*. Knowing precisely when 'causally relevant' values occur would require repeated measurements, if not regular or continuous monitoring, of all time- variant variables — something that is only rarely undertaken, might often be impracticable (and would potentially be impossible for many of the phenomena, characteristics and features involved in disease and healthcare pathways). For these reasons (and for both time-*in*variant and time-*variant* variables), healthcare students, practitioners and researchers must carefully balance whatever information they have regarding: *when* the variables concerned were measured; *how* these measurements were made; and *what* is known about the nature of the phenomena, features and characteristics concerned. Only then might it be possible to confidently assess when — and therefore where, with respect to all other variables included in the DAG — these variables might be most appropriately positioned.

In this regard it is helpful to be able to rely on the fact that measurements made at any given point in time can *only ever* relate to phenomena, features and characteristics that have *already* occurred — i.e. *prior* to or *at* the point in time at which the measurements concerned were made. However, the certainty this provides is often only helpful for those variables that are unlikely to have occurred: over protracted periods of time; and, in particular, during periods of time in which one or more of the *other* relevant variables might have *also* occurred. For this reason, even time- invariant variables may only be amenable to discrete temporal positioning where the period of time over which these might feasibly have occurred is distinct from those of other relevant variables (including other time-*in*variant and time-*variant* variables).

Under these circumstances, determining the temporal sequence of variables considered relevant for inclusion in a DAG often relies far more on *how* these variables were measured than on *when* these measurements were made. Since the meaning of any given variable is ultimately determined by the technique(s) involved in its measurement, and since knowledge of these technique(s) is usually available to those making use of such variables (even if only as encoded in the formatting of data provided second- or third-hand), this information can offer valuable insights on which healthcare students, practitioners and researchers can attempt to infer *when* the phenomena, characteristics or features concerned are most likely to have occurred.

For example, measurements that depend upon *preceding* phenomena (or the *preceding* values of other characteristics and features) make it possible to position the variables concerned in an appropriate temporal sequence with respect to one another, as is the case: with many of the identity-, age-, growth-, development- and exposure-dependent biological and social processes involved in lifecourse-related

disease pathways; and (perhaps) with most (if not all) of the sector-, structure-, resource- and context-dependent operational procedures and processes that exist within healthcare pathways and the delivery of most other services (see, for example, the DAGs in Ellison and De Wet [38]: 259; and Ellison et al. [39]: 3). Likewise, the techniques required to measure time-variant phenomena can often provide substantive clues as to the period of time (or particular point in time) to which any measured value might relate. This is certainly the case with: disease-relevant biological, social and environmental phenomena that vary or change in a more or less coherent fashion over time, and are likely to have been initiated or determined some time before they 'crystallised' in their measured form; and the levels of expertise, efficiency and performance operating within healthcare systems that might only be achieved as a result of decisions made, or systems that evolved and matured, *prior* to their examination and measurement.

These then are the considerations that lie behind two further steps considered helpful for improving the specification of DAGs:

- **Step 6** — identify which of the included *time-invariant* variables are likely to exhibit substantive variation in the precise point in time at which they occur, and carefully (re)position these at the point in time at which (or by which) the phenomena concerned are likely to have occurred for most (or all) of the cases examined (be these study participants or other entities); and
- **Step 7** — carefully consider whether the measured values of *time-variant* variables are likely to have actually occurred or crystallised sometime before their measurements were made and, where necessary, move the variables concerned to an appropriate, earlier position in the sequence of variables included in the DAG.

For the example summarised in Fig. 3 (above), eight of the measured variables were considered time-invariant: DoB; SEX; DM; CRF; IP; INT; SURG; and AKI. Three of these are events (DoB; INT; and AKI) that demarcate discrete periods of the disease and healthcare pathway, and with respect to which all of the other relevant variables can then be positioned. One other variable (SEX) also constitutes a discrete event, but is determined at the point of conception (subject to later confirmation and formal assignment at the point of birth). As such, these four time-invariant variables do not display any relevant temporal variability when estimating the strength, direction and precision of the focal relationship (between INT and AKI) within the context of Escobar et al.'s [25] study.

Of the remaining four time-invariant variables (DM, CRF, IP and SURG), all are determined by aetiological processes, as reflected in subsequent diagnostic assessments (or therapeutic decisions), which might therefore have feasibly occurred (or been made) at somewhat variable points in time between DoB/SEX and INT (where DM, CRF and IP must have occurred) or between INT and AKI (where SURG must have occurred). When generating the DAG summarised in Fig. 3, it was acknowledged that these events might have all preceded the points in time when the clinical assessments and decisions concerned were made. However, it was

assumed that the emergence of DM was most likely to have preceded that of CRF, and that the thromboses on which IP was based were unlikely to have occurred before either DM or CRF — assumptions that might well be at odds with contemporary understanding or empirical evidence concerning the aetiological mechanisms involved.

Unfortunately, Escobar et al.'s [25] article, like the vast majority of published accounts of empirical research, does not (have space to) provide the evidence required to assess (let alone confirm or reject) any of these assumptions. The temporal positioning of these variables was therefore informed as much by speculation (regarding the disease and healthcare pathways thought likely to have been involved) as by the limited clues available from the methods used to measure each variable. Much as these clues (and the modest contribution that the timing of measurements) can help in elucidating the likely temporal sequence of variables considered relevant for inclusion in a DAG; such clues are rarely definitive or incontestable. Drawing DAGs on the basis of temporal logic is therefore as much an 'art' as a 'science', and ultimately depends upon the insight, deliberation, resolute objectivity, temporal perspective and judgement of the healthcare students, practitioners and researchers concerned. And while a temporal approach can add substantial value to the utility of DAGs (not least in the robust test that such DAGs can then provide to more partial professional/operational knowledge and the imperfect paradigms that reflect our incomplete understanding of biomedical and organisational pathways), it often requires diligent reflection, robust debate and a willingness to disclose any substantive uncertainty by offering a range of DAGs describing the alternative temporal sequences considered plausible (see for example [40]; Figure 1: 843).

Meanwhile, of the remaining measured and unmeasured variables included in Fig. 3, four comprise baseline and follow-up measurements of two time-variant variables (HCT_0/HCT_1 and CRT_0/CRT_1), one of which (CRT_1) was integral to (if not essentially synonymous with) the diagnosis of AKI — a variable previously interpreted as time-invariant (given it reflects a diagnostic decision regarding the subsequent health of the patients examined). Indeed, it was for this reason that AKI was considered central to the temporal framework of the study around which all other variables might then be positioned, even though it was measured/diagnosed on the basis of what must have been the known (yet unreported) time-variant follow-up measurement of CRT (hence CRT_1 [25]: 240 PA; 7 AAM). While it is therefore a moot point whether the value of CRT_1 might have *actually* occurred much earlier than the formal diagnosis of AKI, there is little value in considering this possibility since AKI is based upon (and therefore essentially the *same as*) CRT_1. However, it remains possible (and potentially troublesome) that the value of CRT_1 (and its role in the diagnosis of AKI) might have *actually* been achieved *before* HCT_1 or SURG — a possibility with consequences that will be discussed in Sect. 2.5, and at Step 10 (below).

Setting this possibility aside, it nonetheless remains entirely plausible that the values of HCT_0, CRT_0 and HCT_1 may all have been achieved prior to the points at which these were actually measured; and, moreover, that the *measurements* of HCT_0

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 October 2022 doi:10.2

and CRT_0 may have also occurred over a range of different points in time given the study appears to have drawn on data collected somewhat opportunistically during prior clinical care (rather than at predetermined points in time within a carefully monitored disease and healthcare pathway — as might have been the case were the authors to have adopted a more deliberative prospective study design). For this reason, the development of alternative DAGs (as discussed earlier, and in Sect. 2.4 and at Step 9, below) should also consider the possibility that the values of HCT_0 and CRT_0 were measured and/or occurred *before* CRF or DM.

These considerations have been summarised in Fig. 4 in which the potential, alternative temporal positions of the relevant covariates have been highlighted using coloured boxes spanning the periods in time over which these covariates might plausibly have occurred (or 'crystallised'). These reflect substantial uncertainty regarding the *precise* temporal positioning of: DM, CRF, CRT_0 and HCT_0 with respect to one another (although all four covariates are considered likely to have occurred/crystallised *after* DoB/SEX and *before* IP); and SURG, HCT_1 and CRT_1/AKI — both with respect to one another *and* regarding the possibility that the value of CRT_1 (as and when measured, together with the diagnosis of AKI that these values then permitted) may have plausibly occurred/crystallised *before* SURG occurred and/or *before* the value of HCT_1 (as and when measured) had crystallised (i.e. closer to the point at which this was measured, assuming CRT_1 and HCT_1 were measured at the same time).

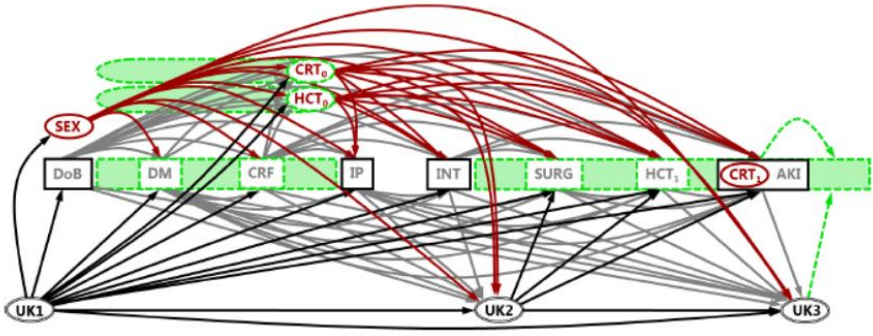


Fig. 4 A further iteration of the DAG presented in Fig. 3 in which the potential, alternative temporal positions of six time-invariant/variant covariates (DM, CRF, CRT_0, HCT_0, SURG, and HCT_1) have been highlighted using boxes (in **green font**) spanning the periods over which these covariates might have plausibly occurred (or crystallised). Were SURG and/or the value of HCT_1 to have occurred/crystallised *after* the value of CRT_1 had crystallised (as and when the values of these variables were measured), additional causal paths might then be required to reflect the plausible probabilistic impact *of* CRT_1/AKI and UK3 *on* SURG and/or HCT_1 (as indicated by the dashed arrows in **green font**). **UK1:** UnKnown/unacknowledged variable set 1; **SEX**: Sex; **DoB**: Date of birth-determined age at intervention; **DM**: Pre-intervention diagnosis of diabetes mellitus; **CRF**: Pre-intervention diagnosis of chronic renal failure; **HCT_0**: Baseline/pre-intervention haematocrit values**;** **CRT_0**: Baseline/pre-intervention creatinine values; **IP**: Indication for the procedure (venous vs. arterial); **UK2:** UnKnown/unacknowledged variable set 2; **INT**: Intervention (PPT vs. CDT); **SURG**: Post-intervention open surgical procedures; **HCT_1**: Post-intervention haematocrit values; **CRT_1**: Post-intervention creatinine values; **AKI**: Acute kidney injury; **UK3**: UnKnown/unacknowledged variable set 3.

2.4 Challenges Facing the Specification of Causal Relationships Amongst Non-asynchronous Variables

Alongside the additional steps required to accommodate omitted and time-variant variables, DAGs also need to be able to deal with subsets of variables that occur or crystallise at (more or less) the same point in time — i.e. those that are 'non-*a*synchronous'. For these groups of variables it is not possible to determine which might act as the cause or consequence of the other(s). It is therefore inappropriate to include causal paths in either direction between non-asynchronous variables within a DAG, even though all such variables are likely to share any preceding (measured, unmeasured and unacknowledged) probabilistic causes and all subsequent probabilistic consequences (notwithstanding definitive evidence to the contrary; see Step 3, above).

Examples of non-asynchronous variables within disease and healthcare pathways might include any number of discrete biological, social or circumstantial phenomena, features and characteristics that emerge as multiple consequences of singular events, such as: the plethora of genotypic characteristics determined at conception; or the metabolic, dietary- and activity-related phenotypes that emerge and crystallise at similar periods of time as a result of the interaction between an individual's genetic heritage, sociocultural background/circumstances, educational opportunities and economic trajectory. Within healthcare pathways, there are similar structural, procedural and contextual characteristics that are likely to: share a common preceding cause (such as dis/investment in resources or substantive changes in policies, contexts, infrastructure, services or facilities); and then lead to simultaneous changes in a wide variety of subsequent phenomena, characteristics and features. When drawing DAGs informed by temporal logic, it is therefore important to: carefully consider whether one or more of the variables included might need to be situated at *essentially* the same point in time within the hypothesised temporal sequence; and take care not to include causal paths between any of these non- asynchronous variables.

However, the decision to position variables at *essentially* the same point in time along the temporal sequence of probabilistic causes and consequences should ordinarily reflect the likelihood that the phenomena, characteristics and features concerned *actually* occurred or crystallised at more or less the same time. Indeed, it would be poor practice to treat variables as non-asynchronous simply whenever there is uncertainty regarding which of these might precede the other. Instead, such uncertainty is best addressed by offering a range of DAGs describing all of the alternative temporal sequences considered plausible (as suggested in Sect. 2.3, above).

These considerations are encompassed in two further steps that can help improve the accuracy and functional utility of DAGs:

- **Step 8** — ensure there are no causal paths between any variables considered non-asynchronous; and

- **Step 9** — wherever substantive uncertainty prevails regarding the temporal position of any variables (following *any* of the preceding steps), provide a comprehensive range of alternative DAGs in which the variables concerned occupy alternative, plausible temporal positions in the sequence of variables included.

For example, when adding four measured yet omitted covariates to the DAG in Fig. 3 (these being: patient sex: SEX; baseline creatinine and haematocrit values: CRT_0 and HCT_0; and follow-up creatinine values: CRT_1), it was judged that the first of these (SEX — a time-*in*variant variable) would have occurred at *essentially* the very same time that each of the patients included in the study were conceived or born (i.e. DoB). In contrast, the second and third of these (CRT_0 and HCT_0 — both of which were time-*variant* variables) were considered likely to have crystallised: over a more diffuse period of time (i.e. as and when measured during the period of prior clinical care encompassed by this study); and at approximately, and therefore *essentially*, the same point in time with regard to one another. This was the rationale behind positioning both of these variables *after* the two time-*in*variant markers of pre-thrombotic chronic disease (diabetes mellitus: DM; and chronic renal failure: CRF) and *before* the onset of thrombosis (whether arterial or venous: IP). Likewise, given the dependence of the specified outcome (AKI) on the (known, measured, yet unreported and initially omitted) follow-up measurement of creatinine (CRT_1), these two variables were considered synonymous and therefore *entirely* (rather than *essentially*) co-terminous.

Were more frequent measures of CRT and HCT available (both before and after the point at which CRT_0 and HCT_0 were *actually* measured), it might have been plausible to select values from amongst these measurements so as to be able to position the two covariates asynchronously within the DAG, with one acting as a plausible probabilistic cause of the other. But from a purely temporal perspective — and in terms of the causal roles these covariates play with regard to the specified exposure variable (INT) and the specified outcome variable (AKI) — the decision to situate CRT_0 and HCT_0 at (*essentially*) the same point in time within the DAG seems entirely appropriate, albeit somewhat pragmatic.

On the basis of these assessments, any alternative DAGs drawn to explore intractable uncertainty regarding the relative position of time-*in*variant variables (particularly those that might have occurred over a substantial period of time), and time-*variant* variables (particularly those whose values, as measured, may have occurred sometime before they were measured), would involve exploring both: the sequence of variables occurring between DoB/SEX and INT (i.e. DM, CRF, CRT_0, HCT_0 and IP); and those occurring between INT and CRT_1/AKI (i.e. SURG, and HCT_1). However, it is important to point out that the number of additional DAGs required to accommodate variation in the temporal positioning of these covariates (5 between DoB/SEX and INT; and 2 between INT and CRT_1/AKI) would be no fewer than 240, and many more were it necessary to accommodate all possible permutations of non-asynchronously occurring pairs or groups of variables therein. As such, Fig. 4 (which seeks to accommodate the plausible, alternative temporal

positions of DM, CRF, CRT_0 and HCT_0); and SURG, HCT_1 and CRT_1/AKI) conceals the very large number of possible permutations (and alternative DAGs) involved.

2.5 Challenges Posed by Intractable Uncertainties Regarding Temporally Obscure Variables

For this reason, although Step 9 can help to accommodate and explore any intractable uncertainty regarding the likely temporal positioning of measured and unmeasured variables, it achieves this only at the cost of generating additional (and potentially very large numbers of different) DAGs. Moreover, the sizeable number of alternative DAGs that are required to accommodate even a modest degree of temporal uncertainty (as was the case with the worked example presented here) is likely to offer little more than pause for reflection wherever these DAGs are intended to support more practical or pragmatic applications (such as: expert or peer review of the theorised pathways they aim to represent; or robust statistical analysis of the underlying data generating mechanisms involved). At the same time, drawing DAGs that depend upon the (potentially sparse) temporal positioning clues available — from whatever techniques were used to measure each of the time-variant and time-invariant variables included — risks relying once more on prior knowledge, insight or experience and (re)introducing the very same cognitive biases that drawing DAGs using temporal logic sought to avoid. Such risks therefore appear inevitable and may only, at best, be subject to a modest degree of mitigation by focussing intently on temporality.

Nevertheless, in most comparable scenarios to that described by Escobar et al. [25], the need to accurately determine the relative temporal position of *every* variable considered relevant to include in any given DAG may not be *strictly* necessary. This is certainly the case for any hypothesised set(s) of unacknowledged variables, whose temporal position(s) are chosen by the healthcare student, practitioner or researcher concerned, and *simply* to reflect where such variable sets might pose potential risks of analytical and inferential bias. This is also the case whenever the intended application of the DAG is solely (or primarily) to acknowledge or reduce the risk of bias when estimating the strength, direction and precision of a *single* causal path (i.e. a single focal relationship) rather than developing a more comprehensive representation of any given data generating mechanism from which the risk of bias might then be addressed for *multiple* causal paths/focal relationships.

Indeed, the benefit of focussing on just one focal relationship when drawing a DAG is evident from the worked example presented in this chapter. This is because the stated aim of Escobar et al.'s study ([25]: 239 PA; 6 AAM) was principally: 'to determine if patients undergoing AJ [AngioJet™] thrombolysis were at increased risk of acute renal injury, when compared with non-AJ CDT [catheter-directed thrombolysis] …'. As such, their study focussed intently on estimating a single focal relationship between one specified exposure variable (INT: the clinical

procedure/intervention allocated) and one specified outcome variable (AKI: acute kidney injury — defined as 'an increase in Cr [creatinine values] >25% of baseline within 72 hours [of the intervention], or an absolute increase of >0.5 mg/dL' [25]: 240 PA; 7 AAM).

With only a single focal relationship involved, the temporal position of all other variables considered relevant to include in the DAG inevitably falls into one of just three discrete periods of time: before the specified exposure variable; after the specified outcome variable; or somewhere between the two. Crucially, it is these three periods (rather than the specific sequence of variables within each of these periods) that determines: whether any of the other variables (i.e. the covariates) might act as plausible probabilistic causes or consequences of the specified exposure and/or outcome variables; and which of these might then pose a risk of conditioning-dependent analytical and inferential bias when estimating the strength, direction and precision of the focal relationship (as discussed in Sect. 3.2, below). For this reason — and whenever the principal aim of generating a DAG is to eliminate, mitigate or simply acknowledge the risk of bias when estimating a single focal relationship — provided it is possible to determine whether each of the covariates occurred or crystallised before, after or in-between the specified exposure and outcome variables concerned, it is rarely necessary to correctly determine the respective temporal position of those covariates occurring *within* any of these three distinct periods of time.

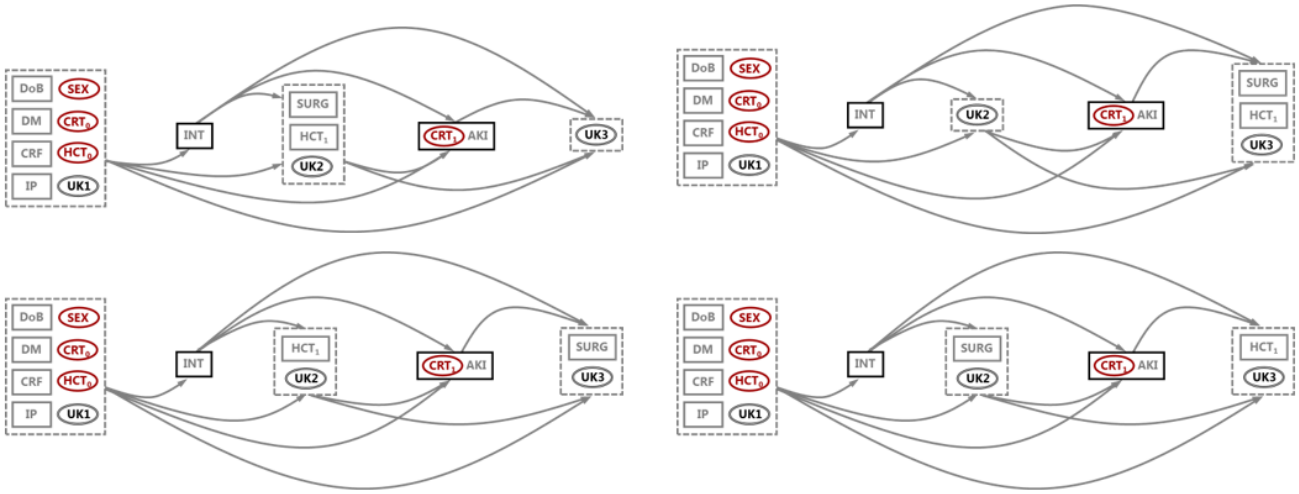
There is therefore a final step that is well worth considering to simplify the representation of causal pathways within a DAG, particularly when these pathways face intractable uncertainty as a result of temporally obscure variables (including those variables likely to have occurred or crystalised over protracted periods of time) — albeit only when the principal aim of the DAG is to explore or address potential sources of analytical and inferential bias when estimating a *single* focal relationship:

- **Step 10** — reduce the number of alternative DAGs generated at Step 9 by creating three discrete covariate 'sets' comprising all those whose temporal position can confidently be located: before the specified exposure variable; after the specified outcome variable; or in-between the two.

To this end, Fig. 5 provides four further iterations of the DAG summarised in Fig. 3 in which each covariate has been included in one of three discrete sets of variables located: before the specified exposure variable (INT); after the specified outcome variable (CRT_1/AKI); or in-between the two. Not only can this approach substantially reduce the number of alternative DAGs required to accommodate uncertainty in the temporal positioning of covariates *within* each of these three periods; but it does so without recourse to aetiological knowledge, or professional/operational experience (and the risk of cognitive bias that reliance thereon can entail). In the case of Escobar et al.'s [25] study, where only two of the available covariates (SURG and HCT_1) were considered likely to have occurred in *more than one* of these three periods of time (see Sect. 2.3 and Figure 4, above), only four alternative DAGs are required to summarise the key causal relationships between

doi:10.2

Figure 5. Four simplifications of the 240+ alternative versions of the DAG presented in Figure 2.1 (redrawn from the DAG provided by Escobar et al., 2017: 244; Figure 7: 244 PA; 29 AAM) required to accommodate residual uncertainty in the temporal positioning of two key covariates (SURG; and HCT_1) in which each of the measured (**grey font**), unmeasured (**red font**) and unacknowledged (**black font**) covariates has been included in one of three discrete sets of variables, enclosed within separate boxes and located either: before the specified exposure variable (INT); after the specified outcome variable (AKI/CRT_1); or in-between the two. This approach to simplifying DAGs clearly reduces the very large number of alternative DAGs required to accommodate uncertainty in the temporal positioning of covariates therein, while retaining their utility for identifying those covariates that present a potential risk of analytical and inferential bias when estimating the strength, direction and precision of the specified focal relationship (in this instance between INT and CRT_1/AKI). See: DAGitty.



Key:

- **UK1:** UnKnown/unacknowledged variable set 1
- **DoB:** Date of birth-determined age at intervention
- **CRF:** Pre-intervention diagnosis of chronic renal failure
- **CRT_0:** Baseline creatinine values
- **UK2:** UnKnown/unacknowledged variable set 2
- **SURG:** Post-intervention open surgical procedures
- **CRT_1:** Post-intervention creatinine values
- **UK3:** UnKnown/unacknowledged variable set 3
- **SEX:** Sex
- **DM:** Pre-intervention diagnosis of diabetes mellitus
- **HCT_0:** Baseline/post-intervention haematocrit values
- **IP:** Indication for the procedure (venous vs. arterial)
- **INT:** Intervention (PPT vs. CDT)
- **HCT_1:** Post-intervention haematocrit values
- **AKI:** Acute kidney injury

each of the covariates and the specified exposure and outcome variables — a substantial simplification of the 240+ possible DAGs considered necessary following Step 9 (above).

3 Using DAGs to Identify/Address Potential Sources of Analytical or Inferential Bias

3.1 Identifying Potential Errors and Biases Associated with Epistemological Assumptions

A key benefit of using DAGs to represent disease and healthcare pathways (and the data generating mechanisms involved therein) is that these representations can help to expose potential errors and biases associated with the epistemological assumptions that all such diagrams require. Indeed, we have already seen how DAGs informed by temporal logic might help to assess whether the causal paths considered likely or necessary on the basis of prevailing aetiological knowledge, and professional/operational experience are consistent with those considered possible or plausible on the basis of temporality and temporal logic. These assessments of internal and external consistency can inform, or be prompted by, decisions made at each and every stage of drawing a DAG. For example, when the variables considered relevant to include are arranged in a temporal-causal sequence, it soon becomes apparent if *none* of the covariates act: as plausible, probabilistic causes of the exposure and/or outcome; or as variables that fall along the causal path between the two. The most helpful DAGs in this regard will have covariates located before, after *and* in-between the specified exposure and outcome variables in order to prompt reflection on the specific roles that variables positioned in each of these locations might perform (and any associated risks of conditioning-dependent bias when estimating the strength, direction and precision of the focal relationship involved — see Sect. 3.2, below). Moreover, the absence or omission of covariates within any of these locations should prompt substantial reflection and discussion amongst the healthcare students, practitioners and researchers concerned regarding what additional variables (whether measured or unmeasured, known or unknown) might be relevant, appropriate, helpful or necessary to include, including those that might as yet be unacknowledged (such as UK1, UK2 and UK3 in Figs. 3, 4 and 5).

3.2 Identifying Potential Sources of Conditioning-Dependent Analytical and Inferential Bias

From an analytical and statistical perspective, an even more important benefit of using DAGs to represent disease and healthcare pathways, and the underlying data generating mechanisms these pathways involve, is that they help to identify potential sources of conditioning-dependent bias that can affect the estimated strength, direction and precision of any specified focal relationship [41, 42]. These biases are the parametric consequences of the temporal-causal relationships that

exist between each of the covariates and the specified exposure and outcome variables. Where these covariates precede both the specified exposure and specified outcome variables they will act as potential 'confounders' (i.e. plausible, probabilistic causes of *both* the exposure and the outcome [41]) and — unless their effects are removed by conditioning on the covariates concerned (whether through selective sampling, stratification or statistical adjustment [43]) — the estimated strength, direction and precision of the specified focal relationship may then be susceptible to 'confounding bias'. In contrast, wherever covariates are plausible, probabilistic *consequences* of both the exposure and the outcome (i.e. where they might be called 'consequences of the outcome' [16]), or where they act as 'mediators' that fall along the speculative causal path between the exposure and outcome (being plausible, probabilistic *consequences* of the exposure, and plausible, probabilistic *causes* of the outcome [44]), the so-called 'collider bias' that these can contribute to the estimated strength, direction and precision of the specified focal relationship will *only* occur following conditioning thereon (whether, as before, through selective sampling, stratification or statistical adjustment [45–48]).

DAGs therefore have substantial utility for: classifying which of the covariates might act as confounders, consequences of the outcome or mediators; and thereby evaluating the potential contribution that each might make to conditioning-dependent analytical and inferential bias when estimating a specified focal relationship. In particular, DAGs help to prompt careful reflection amongst the healthcare students, practitioners and researchers involved regarding: the potential risk of conditioning on consequences of the outcome or mediators as a result of sampling, stratification or statistical adjustment decisions (whether deliberately, unintentionally or as a result of misclassification [43]); and whether the inclusion, measurement and adjustment for alternative confounders (and particularly those offering more accurate and more comprehensive assessments of the potential for confounding) might help further mitigate the risk of residual or unadjusted confounding bias, respectively (see [17, 18]; and [16]).

For example, returning to the DAG presented in Fig. 3 (based on the DAG provided by Escobar et al. [25]: Figure 7: 244 PA; 29 AAM), the authors concerned might have found it insightful to consider whether:

- the estimated relationship between the specified exposure variable (INT) and the specified outcome variable (CRT_1/AKI) might have been subject to collider bias as a result of conditioning on one or more of: the measured mediators (i.e. SURG; and HCT_1); the set of unacknowledged mediators (i.e. one or more covariates included in UK2); or the set of unacknowledged consequences of the outcome (i.e. one or more of the covariates included in UK3);
- it might be worth conditioning on the additional potential confounders (SEX; CRT_0; and HTC_0) — that were measured yet were omitted from (and could have been included in) Escobar et al.'s [25] original DAG — in order to further mitigate the likely effects of unmeasured and unadjusted confounding bias [49];

- it might likewise be worth conditioning on any additional unmeasured potential confounders, including those as yet unacknowledged (i.e. those within UK1) — assuming either that: data on appropriate variables might be available within the medical records of the patients examined; or it would be feasible to repeat the analyses with a subsequent cohort of patients for whom some of these additional covariates might then be identified and measured; and
- any of the covariates situated immediately *before* and immediately *after* the specified exposure variable (such as: IP, CRT_0 and HTC_0; and SURG and HCT_1, respectively) might have *actually* occurred or crystallised later/earlier (i.e. *after* or *before* the specified exposure variable) — resulting in: conditioning on mediators that had been misclassified as potential confounders (in this instance: IP; CRT_0; and/or HTC_0); or a failure to condition on potential confounders that had been misclassified as mediators (in this instance: SURG; and/or HCT_1).

3.3 Mitigating Potential Sources of Bias When Selecting the Datasets and Statistical Adjustment Sets Used

In each of these instances, a DAG (such as that presented in Fig. 3) can clearly help provide a credible, principled and interrogable basis for exploring whether improvements in a number of key methodological decisions might help to mitigate the risk of analytical and inferential bias during the estimation of specified focal relationships — including:

- the sampling or stratification of participants included in a prospective study — to reduce the risk of collider bias as a result of sampling or stratification that deliberately or unintentionally conditions on a mediator or consequence of the outcome;
- the choice of any retrospective or existing dataset(s) selected for secondary analysis — again to reduce the risk of collider bias by carefully ensuring that any datasets selected for analysis are unaffected by sampling or stratification that might have deliberately or unintentionally conditioned on a mediator or a consequence of the outcome;
- the measurement of additional or alternative covariates (either prospectively or by those involved in generating the existing dataset[s] subsequently selected for secondary analysis) — to reduce the risk of inappropriate conditioning on mediators or consequences of the outcome misclassified as confounders; and to mitigate residual or unmeasured (and therefore unadjusted) confounding bias through statistical adjustment of more accurately measured or a more comprehensive selection of potential confounders, respectively; or
- the exclusion of misclassified confounders, and the inclusion of more accurately measured (or a more comprehensive selection of additional or alternative) confounders in the covariate adjustment sets of the multivariable statistical models used to estimate the strength, direction and precision of the specified focal relationship — again to reduce the risk of inappropriate conditioning on

mediators or consequences of the outcome misclassified as confounders; and to mitigate residual or unmeasured (and therefore unadjusted) confounding bias through statistical adjustments of more accurately measured, or a more comprehensive selection of, potential confounders, respectively.

3.4 Identifying Potential Sources of Bias During the Critical Appraisal of Published Analyses

Each of these potential improvements in methodological decisions that DAGs are able to support also lend themselves to the critical appraisal of previously published analyses, regardless of whether these made use of DAGs (and did so competently or otherwise). Indeed, after extracting the variables that were included by Escobar et al. [25] in their DAG (Fig. 7: 244 PA; 29 AAM) and/or Table 1 (241 PA; 21 AAM), and after identifying the specified exposure (INT) and outcome (AKI) variables involved in their study's principal focal relationship, the worked example presented in this chapter illustrates how it can be relatively straightforward to: generate a DAG (even when the authors concerned did not use or report one, as Escobar et al. [25] did) using each of the ten steps described in Sect. 2, above; and classify which of the measured and/or included covariates might act as confounders (in this instance: DoB; Sex; DM; CRF; CRT_0; HCT_0; and IP), potential mediators or consequences of the outcome (in this instance: SURG; and HCT_1) for the focal relationship concerned (see Fig. 5). Based on this information, and after careful consideration of the data sources, sampling techniques, measurement protocols and primary/secondary datasets used, it should also be possible to assess whether any of the estimates generated for the principal focal relationship(s) examined might have been affected by prospective study designs or a reliance on existing, secondary datasets that involved:

- sampling or stratification (whether deliberate or unintentional) that was likely to have conditioned on a mediator or consequence of the outcome — thereby increasing the risk that the estimates generated might have been affected by collider bias;
- inadequate numbers of covariates, or covariates measured with insufficient accuracy or precision, to permit robust conditioning on/adjustment for those classified as confounders — thereby increasing the risk that the estimates generated were affected by unmeasured/unadjusted or residual confounding, respectively; and
- the misclassification of, and inappropriate conditioning on, mediators or consequences of the outcome as if these were confounders — thereby increasing the risk that the estimates generated were inadvertently affected by collider bias as a result of conditioning on these (misclassified) colliders.

4 A Role for Educators in Addressing the Limitations of Directed Acyclic Graphs

4.1 Transparent Reporting of Context-Specific Uncertainties, Assessments and Potential Errors

DAGs can only ever provide imperfect *theoretical* approximations of causal pathways and data generating mechanisms, not least because — except where these pathways and mechanisms are deliberate or incidental human artefacts that have been faithfully rendered or comprehensively described, respectively — it is simply implausible that we will ever know: *all* of the relevant variables involved; *which* of these are necessary or sufficient to include therein; and *where* each of these should be temporally and causally positioned with respect to one another. Indeed, beyond the challenges posed by omitted, time-variant, non-asynchronous and temporally obscure variables, it is also likely that many phenomena, characteristics and features (*both* time-invariant *and* time-variant) occur or crystallise over a sufficiently broad period of time that their causally relevant temporal positions within a DAG overlap with (many) other variables, making it very difficult to determine which (if any) are causes or consequences of each of the others. Wherever the extent of these overlaps makes it unclear whether the variables concerned act as confounders rather than mediators or consequences of the outcome, it will severely limit the practical utility of the DAGs concerned. This is because such DAGs will involve substantial uncertainty as to which of the covariates require conditioning (to mitigate confounding bias), and which do not (to avoid collider bias).

These epistemological uncertainties remind us that DAGs, like all tools, are rarely infallible. Moreover, as principled yet highly conceptual models, DAGs can prove elusive and difficult to articulate and operationalise; and the level of skill, diligence and discipline required to grasp and apply the principles on which DAGs are based can impose very real limitations on their practicability and the value of the evidence they generate. As such — and like any novel tool contingent on the knowledge, understanding and competence required for their successful application — DAGs can somewhat paradoxically: reduce (rather than enhance) the thoughtfulness involved in the design, application and interpretation of inferential statistics; obfuscate or conceal thoughtless and incompetent practices; offer an entirely false aura of novelty, sophistication and expertise; and make it *more* (not *less*) difficult for others to critique and challenge the theories and analyses involved, even when these are patently misguided or flawed [50]. Indeed, the recent review by Tennant et al. [35] found substantial evidence of simplistic, wrong-headed, thoughtless and ostensibly sloppy applications of DAGs, including: a remarkably large proportion of articles (38%) that reported using DAGs but failed to share them (thereby eliminating the possibility of evaluating the theoretical integrity or temporal plausibility of these DAGs); very few (<3%) that included all of the plausible, probabilistic causal paths; and fewer than half that reported either the DAG-implied covariate adjustment set(s) (49%) *or* the results of multivariable statistical models where these adjustment sets had been faithfully applied (43%).

These weaknesses in the application and reporting of DAGs mean that their growing popularity may not necessarily lead to consistent improvements in the elimination, mitigation or acknowledgement of bias resulting from inappropriate conditioning on: one or more colliders; or an inadequate selection of incorrectly classified, or imprecisely measured, confounders. Their use is also unlikely to *successfully* bridge the divide between professional expertise and statistical inference *without*: substantial improvements in practice; and greater clarity and transparency in reporting. To this end, Tennant et al. [35] offered eight succinct recommendations to improve both the application and reporting of DAGs; and offered a useful checklist to assist authors preparing (and reviewers evaluating) articles using DAGs (Supplementary Table S6). These have been summarised in the appendices to this chapter (see Appendix 2), and augmented with two further recommendations for improving the interpretation and interrogability of DAGs, namely: the importance of reporting the intended *application* of the DAG (whether analytical, speculative or hypothetical); and using accessible software (such as DAGitty [51–53]) when drawing and sharing DAGs, to avoid the limitations of static two-dimensional visual representations which can be difficult to read, and particularly so when they contain more than a modest number of variables and causal paths (compare, for example, Figs. 2, 3, 4 and 5).

4.2 A Role for Educators in Optimising the Utility of Directed Acyclic Graphs

The aim of this chapter has been to encourage the adoption, and improve the application, of DAGs by demonstrating how educators might use them to facilitate the acquisition of secondary/ancillary skills in inferential statistics by healthcare students, practitioners and researchers. The rationale for this approach stems from the dual utility of DAGs as principled; qualitative, nonparametric representations of temporal-causal disease and healthcare pathways; and of the underlying data generating mechanisms these involve — i.e. the mechanisms on which the quantitative, parametric insights generated using inferential statistics depend. While DAGs require thoughtfulness and a willingness to accommodate uncertainty, they help to prompt careful reflection and offer far greater transparency regarding the assumptions involved when applying inferential statistics to the analysis of essentially theoretical (and often extensively speculative or entirely hypothetical) disease and healthcare pathways.

These benefits accrue because DAGs have substantial utility for bridging the divide between: the aetiological knowledge and professional/operational experience on which contemporary clinical training and practice depend; and the more abstract epistemological and analytical considerations required to extract robust statistical insight from health and healthcare data. Moreover, provided it is plausible to generate finite and manageable numbers of alternative DAGs that accommodate any uncertainty regarding the temporal positioning of critical covariates (particularly those that might plausibly occur before or after the specified exposure variable),

DAGs can still offer a coherent approach to the assessment and application of alternative conditioning strategies involving sampling, stratification or adjustment; and thereby support bias-mitigated estimation of the strength, direction and precision of the focal relationship(s) examined. Indeed, where these estimates are largely unaffected by the speculative re-positioning of covariates that might plausibly occur or crystallise before vs. after the specified exposure variable (as might occur where the impact of the confounder and collider bias involved is weak), then the use of multiple, alternative DAGs may even provide substantial reassurance that the estimates generated are likely to be robust, and therefore have substantial practical utility.

For these reasons, DAGs represent an invaluable educational tool for helping to improve both the theoretical and analytical capabilities of healthcare students, practitioners and researchers by:

- obliging them to carefully specify any explicit and implicit assumptions that underpin the theoretical causal pathways on which their analytical designs are based;
- enabling them to explore the temporal plausibility, and any associated uncertainty, regarding the role(s) that measured, unmeasured, unmeasurable and unacknowledged covariates might play therein;
- helping them to address any temporal implausibility and uncertainty in the sampling, measurement and analytical assessments and decisions they make; and
- ensuring they can then acknowledge any residual risk of conditioning- dependent analytical and inferential bias when interpreting and reporting their findings.

The step change in capability, transparency and epistemological competency that each of these benefits provide should offer a compelling case for educators to integrate DAGs into the statistical training and support provided to healthcare students, practitioners and researchers. To facilitate this, the worked example summarised in this chapter can be used to generate a lesson plan or self-directed learning exercise covering each of the ten steps for drawing a DAG using temporal logic (as outlined in Sect. 2, above), and to extend the application of each of these steps beyond the critical appraisal of published analyses that focus on the estimation of (one or more) focal relationships, to demonstrate how related considerations might be usefully applied to: inform prospective studies relying on either secondary or primary datasets; and ensure these datasets are selected or generated so as to optimise the opportunities available for eliminating or mitigating analytical and inferential bias when estimating a clearly specified focal relationship.

Acknowledgements Johannes Textor and Mark Gilthorpe kindly shared many of the perspectives and insights (and much of the understanding) required to develop the approach for drawing DAGs presented in this chapter. I would also like to thank Peter Tennant and Laust Mortensen for sharing the approaches they use when explaining the meaning and utility of DAGs; and Tomasina Stacey, Renata Medeiros Mirra and Damian Farnell for their helpful comments and suggestions on previous versions of the chapter.

Appendix 1: A Comprehensive List of Variables Considered Relevant by Escobar et al. [25]

Although the worked example summarised in this chapter focussed only on those variables that Escobar et al. [25] chose to include in their published DAG (Figure 7: 244 PA; or 29 AAM) and Table 1 (241 PA; or 21 AAM), this constitutes a gross simplification of Step 1 ('Compile a comprehensive list of any phenomena, characteristics or features [and their associated variables] considered relevant to the focal relationship.'). Indeed, the simplified approach to Step 1 adopted in this chapter is only likely to be sufficient for critically appraising the temporal-causal plausibility of Escobar et al.'s [25] published DAG itself; and is unlikely to support a substantive assessment of any potential, additional and residual risks of analytical and inferential bias associated with the omission of relevant and important variables from their DAG. This is because whenever a more *comprehensive* assessment is required — of the variables and causal paths included and omitted from DAGs developed to support the estimation of specified focal relationships — it is necessary to generate a more *comprehensive* list of the phenomena, characteristics and features involved (*both* when designing DAG-informed studies/analyses *and* when evaluating their published findings).

In both instances, substantial reflection and discussion amongst the healthcare students, practitioners and researchers concerned can help to carefully explore all of the (known and unknown) phenomena, characteristics and features that might plausibly be relevant to the specified focal relationship(s) involved — and therefore what associated variables ought to be/have been included in the causal path diagram(s)/ DAG(s) used, whether to: generate (or select) primary (or secondary) datasets; or mitigate the risk of bias associated with the sampling or analytical constraints these datasets impose/involve. Substantial reflection and discussion may also be evident in the deliberations and decisions reported in published accounts of such studies, and these can provide a helpful list of any ancillary variables the authors considered relevant to each of the focal relationships and sampling frames involved (and, therefore, any additional variables that might have been useful or necessary to include in each of the studies' DAGs). Moreover — whenever the reviewer(s) involved in critically appraising such studies lacks the knowledge, insight and experience required to generate an extensive (or comprehensive) list of the phenomena, characteristics, features and associated variables likely to be relevant to/implicated in the focal relationship(s) concerned; or is particularly interested in the ideas, insights, views and perspectives provided by the authors involved — collating a detailed list of the variables cited in the study's rationale, design and interpretation (such as those summarised in the article's Introduction and Discussion sections) can prove invaluable.

To illustrate what this can entail when critically appraising published studies using DAG-informed designs and analyses, it is worth compiling a list of all of the phenomena, characteristics, features and associated variables that Escobar et al. [25]

cited when developing, implementing, interpreting and reporting their study. Excluding those required to record the specified exposure and outcome variables, this list of variables includes any mentioned as potentially relevant for consideration in the study's rationale, design and interpretation (such as those summarised in the article's Introduction or Discussion sections), as well as any variables that were *actually* available to, and were measured, summarised and/or analysed in the Methods and Results sections of the article.

Careful and close reading of Escobar et al.'s [25] article identified numerous phenomena, characteristics, features and variables that had been cited as relevant to their principal focal relationship. These included those considered likely to: reflect explicit or implicit inclusion and exclusion criteria (on the basis of which the patients examined were selected for analysis); determine whether the patients concerned received each of the two clinical procedures/interventions (PPT vs. CDT); mediate the impact of these clinical interventions on the specified outcome (AKI); and emerge as immediate or longer term sequelae of the clinical procedure/intervention received (and/or the specified outcome, AKI itself).

Excluding the specified exposure and outcome variables, covariates relevant to each of these phenomena, characteristics and features were cited in: every section of the article (including the Introduction, Methods, Results and Discussion sections); both Tables; and all seven Figures, as summarised below:

Introduction: 'acute arterial thrombotic syndrome'; 'acute venous thrombotic syndrome'; 'treatment time [duration]'; 'total volume of fibrinolytics'; 'intensive care unit length of stay'; 'hospital length of stay'; 'hospital costs'; 'bleeding complications'; 'hemolysis'; 'hematuria'; 'renal failure'; 'renal function'; 'dialysis'.

Methods: 'patients at our institution'; 'from 2007 to 2013 [hence date on which treatment was provided]'; 'The approach of treatment was up to the practitioner at the time of the [procedure/]intervention, and not protocoled [hence attending physician/practitioner]'; 'demographics, indications, laboratory values before and after the procedure[/intervention] (up to 3 days) [see below for the specific variables involved]'; 'chronic kidney disease'; 'any patient with that diagnosis documented by a nephrologist before our procedure[/intervention — hence documenting nephrologist]'; 'baseline creatinine (Cr)'; 'contrast-induced nephropathy (CIN)'; 'increase in Cr [creatinine] within 72 hours [of the procedure/intervention]'; 'patients on dialysis before AJ'; '[patients with] duplicated codes'; '[patients] without laboratory values obtained before and 24–72 hours after treatment[/procedure/intervention]'; 'hematocrit (HCT)'; 'glomerular filtration rate (GFR)'; '[patient] age'; '[patient] weight'; 'expos[ure] to 270 mgI/ mL Iodixanol … a hypo-osmolar, non-ionic iodinated contrast agent'.

*Results: '*Cr kinase'; 'myoglobin'; '[patient] age'; 'baseline Cr'; 'HCT before treatment'; 'arterial thrombolysis … indication [for the procedure/intervention]'; 'venous [thrombolysis … indication for the procedure/intervention]'; 'rise in Cr'; 'preoperative chronic renal failure (CRF)'; 'dialysis within 2 days of …

procedure/ intervention'; 'died [mortality] during hospitalization'; 'mesenteric ischemia'; 'pulmonary embolus'; 'open surgery'; 'thromboembolectomy'; 'fasciotomy[y]';

'bypass or endarterectomy'; 'major amputation'; '[patient sex] male [vs. female]'; 'diabetes mellitus (DM)'; 'postprocedure[/intervention] HCT'; 'blood loss [as a synonym for change in HCT?] … percentage change from baseline'.

Table I: '[patient] age'; 'male [patient sex]'; 'diabetes'; '[baseline] Cr'; 'arterial indication [for procedure/intervention]'; 'venous indication [for procedure/intervention]'; '[baseline] HCT'.

Table II: 'angioplasty [endovascular procedure/intervention in addition to AJ or CDT]'; 'stenting [endovascular procedure/intervention in addition to AJ or CDT]'; 'fasciotomy [post-procedure/intervention open surgical procedure/intervention]'; 'thromboembolectomy [post-procedure/intervention open surgical procedure/intervention]'; 'bypass/endarterectomy [post-procedure/intervention open surgical procedure/intervention]'; 'major amputation [post-procedure/ intervention open surgical procedure/intervention]'.

Figure 1: 'hematuria'.

Figure 2: 'Cr [change from baseline post-procedure/intervention]'.

Figure 3: 'change in creatinine [from baseline post-procedure/intervention]'; 'open surgery [post-procedure/intervention]'.

Figure 4: 'change in hematocrit [from baseline post-procedure/intervention]'.

Figure 5: 'maj surg … major surgery [major open surgery procedures/interventions]'; '[patient] age'; 'male … [patient sex]'; 'diabetes'; 'CRF … chronic renal failure'; 'drop in HCT … hematocrit [from baseline post-procedure/ intervention]'.

Figure 6: 'maj surg … major surgery [major open surgery procedures/interventions]'; '[patient] age'; 'male … [patient sex]'; 'diabetes'; 'CRF … chronic renal failure'; 'drop in HCT … hematocrit [from baseline post-procedure/intervention]'.

Figure 7: 'venous clot [indication for procedure/intervention]'; 'arterial clot [indication for procedure/intervention]'; 'CRF [chronic renal failure]'; '[patient] age'; 'DM [diabetes mellitus]'; 'drop_hematocrit' [from baseline post-procedure/ intervention]'; '[open] surgery [post-procedure/intervention]'.

Discussion: 'as young as … [patient age]'; 'diabetes'; 'final HCT [post-procedure/ intervention]'; 'change in HCT … drop in HCT [from baseline]'; '[open] surgical procedures[/interventions]'; 'anemia'; 'transfusion'; 'iatrogenic [hemo]dilution'; 'Cr … fallen … change [from baseline]'; 'hemolysis'; 'renal failure'; 'free serum haemoglobin'; 'plasma haptoglobin'; 'heme'; 'globin'; 'Tamm-Horsfall proteins'; 'intra[renal]tubular casts'; '[renal] tubular flow'; 'oliguria'; 'azotemia'; 'venous occlusions'; 'renal function'; 'thromboembolic complications after angioplasty'; 'hemoglobinuria'; 'renal dysfunction'; 'patients … who died [mortality]'; 'limb ischemia'; '[renal failure requiring] dialysis'; 'dialysis dependence'; 'acute [renal] tubular necrosis'; 'overwhelming intravascular hemolysis'; 'fulminant renal failure'; 'hemosiderin deposit[s] … in the kidneys'; 'intracranial hemorrhage'; 'arterial blood flow'; 'CIN [contrast-

induced nephropathy]'; '270 mgI/mL Iodixanol'; 'time [duration] of treatment'; 'number of contrast boluses'; 'rhabdomyolysis from ischemia/reperfusion'; 'myoglobin [at baseline]'; 'myoglobin [post-procedure/intervention]'; 'Cr kinase [at baseline]'; 'Cr kinase [post-procedure/intervention]'; 'amputations'; 'fasciotomies'; 'venous thrombolysis [indication for procedure/intervention]'; 'arterial [thrombolysis indication for procedure/intervention]'; 'recurrent thrombotic syndromes'; 'time intervals between treatments [/procedures/interventions]'; 'duration of hematuria [post-procedure/intervention]'; 'blood loss'.

In total, perhaps as many as 75 discrete sampling criteria and covariates were cited within Escobar et al. ([25]; see Table 1), although a good number of these: may have been functionally synonymous (such as: 'renal failure' and 'renal (dys)function'); comprised categories or components of other covariates (such as: 'fasciotomy', 'thromboembolectomy', 'bypass/endarterectomy' and 'major amputation', which all appeared to have been classified as 'open surgical procedures'); or constituted derived variables based on repeated measures of one or more other covariates (such as: 'change in creatinine' or 'drop in HCT'). As such, the total number of separate sampling criteria and covariates is likely to be somewhat lower than the 75 listed in Table 1. Nonetheless, it is clear that a sizeable proportion of these (45/75; 60.0%) were only mentioned in just one section of the published article, and the number cited in the Methods and/or Results sections (35/75; 46.7%) may offer a fairer representation of the known sampling criteria and covariates for which measurements were actually available.

Table 1. Variables cited by Escobar et al. [25] with reference to study rationale, design and/or analysis.

Phenomena, characteristics, features and associated variables mentioned/cited	Introduction	Methods	Results	Discussion	Table I	Table II	Figure 1	Figure 2	Figure 3	Figure 4	Figure 5	Figure 6	Figure 7
Acute arterial thrombotic syndrome	x		x	x	x								x
Acute venous thrombotic syndrome	x		x	x	x								x
Treatment duration	x			x									
Total volume of fibrinolytics	x												
Length of stay in intensive care unit	x												
Length of stay in hospital	x												
Hospital costs	x												
Bleeding complications	x												
Hemolysis	x			x									
Hematuria	x						x						
Renal (dys)function	x			x									
Renal failure	x			x									
Dialysis post-procedure	x		x	x									
Patients treated "at our institution"		x											
Date of treatment (2007-2013)		x											
Treatment approach of attending practitioner		x											
Chronic kidney disease		x											
Documenting nephrologist		x											
Creatinine (Cr) at baseline		x	x		x								
Contrast-induced nephropathy (CIN)		x		x									
Change in Cr within 72 hr of procedure		x	x	x				x	x				
Dialysis prior to procedure		x											
Patients with duplicated codes		x											
Patients with missing laboratory values		x											
Hematocrit (HCT) at baseline		x	x		x								
Glomerular filtration rate (GFR)		x											
Age of patient		x	x	x	x						x	x	x
Weight of patient		x											
Exposure to 270 mgI/mL Iodixanol		x		x									
Cr kinase at baseline			x	x									
Cr kinase post-procedure			x	x									
Myoglobin at baseline			x	x									
Myoglobin post-procedure			x	x									
Chronic renal failure at baseline			x								x	x	x
Mortality during hospitalization			x	x									
Mesenteric ischemia			x										
Pulmonary embolus			x										
Open surgery post-procedure (generic)			x	x					x		x	x	x
Thromboembolectomy			x			x							
Fasciotomy			x	x		x							
Bypass or endarterectomy			x			x							
Major amputation			x	x		x							
Sex of patient			x		x						x	x	
Diabetes mellitis (DM) at baseline			x	x	x						x	x	x

Change in HCT within 72 hr of procedure			x	x						x	x	x	x
Angioplasty				x		x							
Stenting						x							
HCT post-procedure				x									
Anemia				x									
Transfusion				x									
Iatrogenic hemodilution				x									
Free serum haemoglobin				x									
Plasma haptoglobin				x									
Heme				x									
Globin				x									
Tamm-Horsfall proteins				x									
Intratubular casts				x									
Renal tubular flow				x									
Oliguria				x									
Azotemia				x									
Venous occlusions				x									
Thromboembolic complications post-angioplasty				x									
Hemoglobinuria				x									
Limb ischemia				x									
Dialysis dependence				x									
Acute renal tubular necrosis				x									
Renal hemosiderin deposits				x									
Intracranial hemorrhage				x									
Arterial blood flow				x									
Number of contrast boluses				x									
Rhabdomyolysis from ischemia/reperfusion				x									
Recurrent thrombotic syndromes				x									
Time intervals between repeat procedures				x									
Duration of hematuria post-procedure				x									
Blood loss post-procedure				x									

Appendix 2: Ten Recommendations to Improve the Application and Reporting of DAGs

The recommendations summarised below include eight proposed by Tennant et al. ([35]: 628–629; Supplementary Table S6) and two additional recommendations intended to improve the interpretation, presentation and interrogability of DAGs. The first eight of these recommend that:

1. *The focal relationship(s) examined (i.e. the relationship[s] between the specified exposure and outcome variables) should always be reported* — a critical consideration given the specification of the focal relationship(s) determines which of the covariates considered relevant for inclusion in the DAG are likely to operate as potential confounders, mediators or consequences of the outcome.

2. *The DAG(s) generated to inform the estimation of each focal relationship should always be reported* — again, a critical consideration given that these DAGs will indicate which of the covariates considered relevant for inclusion therein were/ should have been interpreted as operating as potential confounders, mediators or consequences of the outcome.

3. *Each of the DAGs provided should include all variables considered relevant to the causal pathway(s) and data generating mechanism(s) concerned, including those covariates where no measurements were available (and those that might otherwise remain unacknowledged)* — which will ensure that the authors have explicitly acknowledged the risk that any such covariates might have played as potential sources of collider bias during sampling or stratification, or as sources of unmeasured and unadjusted confounding.

4. *Every DAG should contain all of the plausible, probabilistic causal paths between each of the successive variables considered relevant to include therein, except where there is definitive evidence that any such path(s) cannot or do not exist (in which instance, such evidence should then be reported)* — so that the strong assumption invoked by the omission of (possible and plausible) probabilistic paths is not made without substantive justification; and it is possible to assess any rationale for omitting ostensibly plausible, probabilistic causal paths.

5. *The temporal-causal sequence of measured (and of unmeasured and unacknowledged) variables considered relevant to include in each DAG is best presented in a consistent or coherent direction (whether horizontally, vertically or diagonally)* — so that it is easier to ensure that the DAG is both (uni)*directional* and (strictly) *acyclic*.

6. *The DAG-implied covariate adjustment sets of any multivariable statistical models used to estimate the strength, direction and precision of specified focal relationships should be reported* — so that it is possible to ensure that these adjustment sets include: all measured covariates acting as potential confounders within the DAG; and exclude any covariates acting as mediators or consequences of the outcome, while transparently reporting and acknowledging any risk of unmeasured and unadjusted confounding from unmeasured and unacknowledged covariates acting as potential confounders.

7. *The estimated strength, direction and precision of each focal relationship generated using the precise covariate adjustment set as implied by the relevant DAG should be reported (or, at least, the most complete set available that contains all of the measured covariates likely to act as potential confounders)* — to ensure the insight generated by using a DAG to inform such analyses is reported for consideration and assessment.

8. *The use of alternative, modified or augmented covariate adjustment sets be explicitly justified, such as those containing: only a subset of the measured covariates acting as potential confounders within the DAG; or additional covariates omitted from the DAG, or not considered to be acting as potential confounders therein* — thereby ensuring that the estimates produced when using such adjustment sets can be appropriately interpreted and acknowledged as *potentially* biased.

While the first five of Tennant et al.'s [35] recommendations should all help healthcare students, practitioners and researchers to draw, report and share their DAGs in a consistent, comprehensive and effective fashion, the last three will only be relevant to those studies where DAGs have been used to represent real-world causal processes and thereby inform the selection of covariate adjustment sets to mitigate the risk of confounding bias when estimating the strength, direction and precision of a specified focal relationship using multivariable statistical analyses. Indeed, there are two additional recommendations that might also have substantial relevance: to those who are unaware of the alternative applications of DAGs (i.e. as representations of entirely speculative or hypothetical pathways/mechanisms as opposed to analytical tools used to support inferential statistics [54]); or within analytical contexts and applications where the DAGs involved contain more than a handful of variables, and therefore suffer a substantial decline in readability when generated or shared simply as static, two-dimensional diagrams (see for example: Fig. 3; and [16]).

9. *Since DAGs can be developed to support analytical, speculative or hypothetical applications, students, practitioners and researchers using DAGs should report the intended application(s) for which their DAGs were designed, and the rationale involved when constructing these* — this is because the intended

application(s) and the rationale involved will be central to evaluating the likely value, insight and inference that might be drawn from the design of studies and statistical models based on such DAGs.

10. *Any DAGs in which the number of variables and causal paths included make it unclear whether any of these are present or missing should be shared in interrogable formats that make it possible to ascertain the presence or absence of covariates and impossible vs. plausible/probabilistic causal paths therein* — since, in the absence of such formats, the DAG will provide neither an accurate nor a definitive summary of the variables and causal paths the authors intended to include. Examples of suitable formats include those generated using: WinBUGS [55]; Gephi [56]; or DAGitty [52].

These two recommendations have been applied to the worked example presented in this chapter. First, by making it clear that the DAGs drawn in Figs. 2, 3 and 5 were intended to represent real-world causal processes — i.e. the disease and healthcare pathways relevant to the diagnostic and treatment-related decisions examined by Escobar et al. [25] — and thereby support estimation of the focal relationship between INT and AKI. Second, by providing interrogable versions of these three DAGs redrawn using DAGitty software and published online therein (see: Fig. 2: http://dagitty.net/mp8lf8Z; Fig. 3: http://dagitty.net/mXS3gft); Fig. 5: http://dagitty.net/mjHTJJP — all of which were drawn, and are best viewed, in the 'SEM-like' option available under DAGitty's 'Diagram Style' tab).

References

1. Milic NM, Masic S, Milin-Lazovic J, Trajkovic G, Bukumiric Z, Savic M, Milic NV, Cirkovic A, Gajic M, Kostic M, Ilic A, Stanisavljevic D. The importance of medical students' attitudes regarding cognitive competence for teaching applied statistics: multi-site study and meta-analysis. PLoS One. 2016;11:e0164439, (1–13).
2. Hibbard JH, Peters E. Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. Annu Rev Public Health. 2003;24:413–33.
3. Sinatra GM, Kienhues D, Hofer BK. Addressing challenges to public understanding of science: epistemic cognition, motivated reasoning, and conceptual change. Educ Psychol. 2014;49:123–38.
4. Morris RL. Increasing specialization: why we need to make mathematics more accessible. Soc Epistemol. 2021;35:37–47.
5. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci. 2012;7:1–7.
6. Watt H. Statistics education beyond "significance": novel plain English interpretations to deepen understanding of statistics and to steer away from misinterpretations. In: Medeiros Mirra RJ, Farnell D, editors. Teaching biostatistics in medicine and allied health sciences. Berlin: Springer; 2022.
7. Chapman SJ, Grossman RC, FitzPatrick ME, Brady RR. Randomized controlled trial of plain English and visual abstracts for disseminating surgical research via social media. J Br Surg. 2019;106:1611–6.

8. Larson D, Mazur K, White D, Yarnall C. The user's guide project: looking back and looking forward. J Human Math. 2020;10:411–30.
9. Bartlett G, Gagnon J. Physicians and knowledge translation of statistics: mind the gap. Can Med Assoc J. 2016;188:11–2.
10. Kennedy H, Hill RL, Aiello G, Allen W. The work that visualisation conventions do. Inf Commun Soc. 2016;19:715–35.
11. Passera S. Flowcharts, swimlanes, and timelines: alternatives to prose in communicating legal–bureaucratic instructions to civil servants. J Bus Tech Commun. 2018;32:229–72.
12. Carnap R. Theory and prediction in science. Science. 1946;104:520–1.
13. Tromovitch P. The lay public's misinterpretation of the meaning of 'significant': a call for simple yet significant changes in scientific reporting. J Res Pract. 2015;11:1–11.
14. Castell S, Cameron D, Ginnis S, Gottfried G, Maguire K. Public views of machine learning. London: Ipsos MORI Social Research Institute, Ipsos MORI; 2017. p. 1–92.
15. Hanel PH, Mehler DM. Beyond reporting statistical significance: identifying informative effect sizes to improve scientific communication. Public Underst Sci. 2019;28:468–85.
16. Ellison GTH. Might temporal logic improve the specification of directed acyclic graphs (DAGs)? J Stat Data Sci Educ. 2021;29:202–13.
17. Ellison GTH, Harrison W, Law GR, Textor J. Graphical, cross-tabulatory and relational: which is best for drawing DAGs? Proc VI Eur Congr Methodol. 2014a;6:11.
18. Ellison GTH, Harrison W, Law GR, Textor J. Teaching DAGs to support MBChB students design, analyze and critically appraise clinical research. Proc 35th Annu Burwalls Meet Teach Med Stat. 2014b;35:3.
19. Ben-Shlomo Y, Fallon U, Sterne J, Brookes S. Do medical students with A-level mathematics have a better understanding of the principles behind evidence-based medicine? Med Teach. 2004;26:731–3.
20. Murray S, Gal I. Preparing for diversity in statistics literacy: institutional and educational implications. Proc Sixth Int Conf Teach Stat. 2002;6:1–8.
21. Thompson R, Wylie J, Mulhern G, Hanna D. Predictors of numeracy performance in undergraduate psychology, nursing and medical students. Learn Individ Differ. 2015;43:132–9.
22. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, Miller RA, Digles D, Lopes EN, Ehrhart F, Dupuis LJ, Winckers LA, Coort SL, Willighagen EL, Evelo CT, Pico AR, Kutmon M. WikiPathways: connecting communities. Nucleic Acids Res. 2021;49:D613–21.
23. Manktelow M, Iftikhar A, Bucholc M, McCann M, O'Kane M. Clinical and operational insights from data-driven care pathway mapping: a systematic review. BMC Med Inform Decis Mak. 2022;22:1–22.
24. Young T, Morton A, Soorapanth S. Systems, design and value-for-money in the NHS: mission impossible? Fut Healthc J. 2018;5:156–9.
25. Escobar GA, Burks D, Abate MR, Faramawi MF, Ali AT, Lyons LC, Moursi MM, Smeds MR. Risk of acute kidney injury after percutaneous pharmacomechanical thrombectomy using AngioJet in venous and arterial thrombosis. Ann Vasc Surg. 2017;42:238–45. See: published article [PA]; and author accepted manuscript [AAM].
26. Ensmenger N. The multiple meanings of a flowchart. Inform Cult. 2016;51:321–51.
27. Law GR, Green R, Ellison GTH. Confounding and causal path diagrams. Chapter 1. In: Tu YK, Greenwood DG, editors. Modern methods for epidemiology. Dordrecht: Springer; 2012. p. 1–13.
28. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc. 2016;9:211–7.
29. MacLure M, Schneeweiss S. Causation of bias: the episcope. Epidemiology. 2001;12:114–22.
30. Etchells E. Anchoring bias with critical implications. AHRQ Morbidity and Mortality Rounds on the Web; 2015, p 1–26. https://psnet.ahrq.gov/sites/default/files/import/webmm.ahrq.gov.350_slideshow.ppt.

31. Rehana RW, Huda N. A common heuristic in medicine: anchoring. Ann Med Health Sci Res. 2021;11:1461–3.
32. Ly DP. The influence of the availability heuristic on physicians in the emergency department. Ann Emerg Med. 2021;78:650–7.
33. Redelmeier DA, Ng K. Approach to making the availability heuristic less available. Br Med J Qual Saf. 2020;29:528–30.
34. Aneshensel CS. Theory-based data analysis for the social sciences. London: SAGE Publications; 2002.
35. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, Harrison WJ, Keeble C, Ranker LR, Textor J, Tomova GD, Gilthorpe MS, Ellison GTH. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021;50:620–32.
36. Geer D. Unknowable unknowns. IEEE Secur Priv. 2019;17:80–79.
37. Molony DA. Cognitive bias and the creation and translation of evidence into clinical practice. Adv Chronic Kidney Dis. 2016;23:346–50.
38. Ellison GTH, De Wet T. Johannesburg's 'poor housing, good health' paradox: the role of health status assessment, statistical modelling, residential context and migrant status. Public Health. 2020;186:257–64.
39. Ellison GTH, Mattes RB, Rhoma H, De Wet T. Economic vulnerability and poor service delivery made it more difficult for shack-dwellers to comply with COVID-19 restrictions. S Afr J Sci. 2022;118:1–5.
40. Fleischer NL, Roux AD. Using directed acyclic graphs to guide analyses of neighbourhood health effects: an introduction. J Epidemiol Community Health. 2008;62:842–6.
41. Tennant PWG, Arnold K, Berrie L, Ellison GTH, Gilthorpe MS. Advanced modelling strategies: challenges and pitfalls in robust causal inference with observational data. Leeds: Leeds Institute for Data Analytics (LIDA); 2017. ISBN: 978-1-5272-1208-4.
42. van Zwieten A, Tennant PW, Kelly-Irving M, Blyth FM, Teixeira-Pinto A, Khalatbari-Soltani S. Avoiding overadjustment bias in social epidemiology through appropriate covariate selection: a primer. J Clin Epidemiol. 2022;149:127–36.
43. Blair A, Stewart P, Lubin JH, Forastiere F. Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. Am J Ind Med. 2007;50:199–207.
44. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol. 2013;42:1511–9.
45. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, Poole C. Illustrating bias due to conditioning on a collider. Int J Epidemiol. 2010;2010(39):417–20.
46. Elwert F, Winship C. Endogenous selection bias: the problem of conditioning on a collider variable. Annu Rev Sociol. 2014;40:31.
47. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47:226–35.
48. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009;20:488–95.
49. VanderWeele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. Epidemiology. 2011;22:42–52.
50. Krieger N, Davey Smith G. The tale wagged by the DAG: broadening the scope of causal inference and explanation for epidemiology. Int J Epidemiol. 2016;45:1787–808.
51. Ankan A, Wortel IM, Textor J. Testing graphical causal models using the R package "dagitty". Curr Protoc. 2021;1:e45. https://doi.org/10.1002/cpz1.45, (1–22).
52. Textor J, Hardt J, Knüppel S. DAGitty: a graphical tool for analyzing causal diagrams. Epidemiology. 2011;22:745.
53. Textor J, van der Zander B, Gilthorpe MS, Liśkiewicz M, Ellison GTH. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. Int J Epidemiol. 2016;45:1887–94.

54. Ellison GTH. The strengths and weaknesses of directed acyclic graphs (DAGs) as cognitive, analytical and educational tools for medical statistics. preprints.org. 2022;2 Sep:1–17. https://doi.org/10.20944/preprints202210.0084.v1.
55. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput. 2000;10:325–37.
56. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. Proc Int AAAI Conf Web Soc Media. 2009;3:361–2.