



Published in final edited form as:

Int J Epidemiol. 2025 October 14; 54(6): . doi:10.1093/ije/dyaf207.

Making DAGs Even More Useful: Using Augmented Causal Diagrams to Depict Counterfactual, Study Design, Measurement, Analytical, and Interventional Features

Onyebuchi A. Arah^{1,2,3,4,*}, Matthew M. Coates^{1,2}

¹Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles (UCLA), Los Angeles, CA, USA

²Practical Causal Inference Lab, UCLA, Los Angeles, CA, USA

³Department of Statistics & Data Science, UCLA, Los Angeles, CA, USA

⁴Research Unit for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark

Keywords

directed acyclic graph; augmented directed acyclic graph; causal diagram; graphical augmentation; disease risk score; propensity score

Introduction

Since their mainstream introduction in the 1990s, causal diagrams, including directed acyclic graphs (DAGs), are increasingly used to depict our causal knowledge of the world we study and to guide study design, analysis, and interpretation for causal inference [1, 2]. Beyond describing the data-generating mechanism, DAGs are typically used to select variables for confounding control [2]. However, to do more, researchers have had to modify or augment their DAGs with additional features reflecting theoretical what-if scenarios or study-dependent processes affecting the data, analysis, and interpretation. In this journal, Mansournia *et al.* [3] show how to depict balancing scores on DAGs, a welcome addition to the growing use of augmented graphs. This commentary will first place this work in the broader context of how researchers have tried to do more with augmented causal diagrams, including augmented directed acyclic graphs (ADAGs) which we use more expansively to include all such graphs in this commentary. Then, we highlight some useful features of

*Corresponding author. Department of Epidemiology, UCLA Fielding School of Public Health, 650 Charles E. Young Drive South, Los Angeles, CA 90095, USA. arah@ucla.edu.

Author contributions

OAA drafted the commentary and will act as its guarantor. OAA and MMC contributed to this work and approved the final submission.

Ethics approval

None needed for this commentary since no data were analyzed.

Use of artificial intelligence (AI) tools

No AI tools were used in or for this commentary.

ADAGs that depict balancing scores, focusing on the propensity score for illustration. We conclude with some remarks on the future of graphical augmentation.

Making DAGs even more useful with augmentation

Over the years, DAGs have been augmented to depict counterfactual variables and study design, measurement, analytical, and interventional features [1, 4–10], yielding what we call augmented DAGs (ADAGs). Consider the nonaugmented DAG in Figure 1a with an exposure A , an outcome Y , a mediator M , and two confounders $L1$ and $L2$, which could be used to study the effect of A on Y and the joint effect of A and $L1$ and Y . Suppose we are interested in visualizing non-differential independent measurement error in A , effect-measurement modification of the effect of A on Y by $L1$, interaction between A and M in causing Y , that the confounder $L2$ is unmeasured but its proxies $E2$ and $O2$ are observed, and the consequences of selection on unmeasured $L2$ which affects whether Y is observed.

These study-dependent processes or features can be added to DAG 1a to obtain the ADAG in Figure 1b. For example, the measurement error in exposure is added as a new node A^* with an arrow from A [4, 5]. The effect-measure modification of $A \rightarrow Y$ by $L1$ is visualized using a product node $AL1$ with arrows from A and $L1$ and an outgoing arrow into Y [6]. Similar augmentation draws the interaction between A and M in affecting Y . The selection mechanism that results from selection on $S = 1$ and how S defines whose Y is observed in the study are shown, respectively, with an arrow from unobserved $L2$ (a parent of Y) into S and arrows from S and unobserved Y into observed Y^{obs} in the ADAG 1b [7, 8]. A signed ADAG depicting the direction of effects with + and – symbols can be used for reasoning about uncontrolled confounding (due to $L2$ here) (Figure 1b). Others, such as single-world intervention graphs (as in ADAG 1c) and twin networks (as in ADAG 1d), have been developed to show interventions and thus potential outcomes [chapter 41 in reference 1]. Although standard d-separation can be used to read some key independencies (e.g., conditional exchangeability) for using the PS in these ADAGs to estimate $A \rightarrow Y$, other independence implications of deterministic augmentations need the (modified) D-separation criterion [9].

Doing more with ADAGs that depict balancing scores

Given the growing use of ADAGs, it is unsurprising to see ones depicting balancing scores, namely propensity scores (PS) and disease risk scores (DRS), used in confounding control [3]. Like others [10], Mansournia *et al.* [3] propose inserting the PS between the exposure A and the confounders $L1$ and $L2$, and the DRS between the confounders and the (potential) outcome. [By the way, we believe Mansournia *et al.*'s arrow from A to the computational PS in their Figure 8c should be omitted.] Is it possible to anticipate or design the theoretical balancing scores from the unmodified DAG, like the one in Figure 1e? Is there a general principle for augmenting DAGs with the balancing scores? Which variables should be included? Can we evaluate variable balance using the resulting ADAGs? What else can we infer from them?

Consider the DAG in Figure 1e, adapted from Mansournia *et al.* by including $L0$ (which affects exposure A directly and Y only through A), $L4$ (which affects Y only), and a collider $L3$ (which forms an M -bias structure between A and Y). Under Pearl's Structural Causal Model [1] that would have given rise to DAG 1e, A is given by the function $f_A(L0, L1, L2, U_1, U_A)$ with U_A reflecting the unknown parents of A that are independent of the unknown parents of the other variables in DAG 1e. Then, imagine inventing the PS from DAG 1e by partitioning A into two independent parts which would have arrows from them into A : a deterministic part we call the PS which is formed by the L s in f_A (and represented with arrows from the L s to the PS in ADAGs 1f to 1j), and a variable 'error' part containing at least U_A . Therefore, a general principle for representing the PS as a function of selected L s on the ADAG involves moving any existing d-connection between the L s and A from L s to PS and an arrow from PS to A , as shown in ADAGs 1f to 1j. Nodes like $L3$ or $L4$, which do not cause A but are used to specify PS, can be connected using dashed arrows to distinguish their non-causal role in the partitioning of A . We could conduct a similar visual thought experiment to 'invent' the DRS on an ADAG using the structural model for Y , $f_Y(A, L1, L2, L4, U_2, U_Y)$.

If the L s used in forming the PS suffice to close the backdoor between A and Y , then the PS could also close the same backdoor (see ADAGs 1f to 1j). This answers the question of which variables to include in the PS: at a minimum, those L s ($L1$ and $L2$) that would suffice to 'deconfound' $A \rightarrow Y$. Mansournia *et al.* use a SWIG (a causal diagram augmented with interventional features) to illustrate how conditioning on the DRS satisfies the traditional conditional ignorability/exchangeability assumptions, which could also be shown for the PS. Table 1 presents five PS models using various combinations of $L0$ through $L4$, corresponding to ADAGS 1f to 1j. Only the first four PS models that exclude the collider $L3$ can be expected to result in no open backdoor between A and Y . From Table 1, we also see that a PS model can only be used to achieve variable balance (or independence) between any L and A if that L is included in the PS. Any L not included in a PS (e.g., $L0$ or $L4$ for PS1 in ADAG 1f) retains its original (im)balance across levels of A . Introducing a collider $L3$ into a PS induces collider-bias (ADAG 1j and PS5 in Table 1). ADAGs also give intuition about implications of variable selection for statistical efficiency based on the relationship of selected variables to the exposure and outcome (Table 1).

Finally, from the ADAGs, we can also infer that we could regress Y on the residuals (U_A) from the exposure model ($A = \text{PS} + U_A$) used to create the PS to get unconfounded causal estimates. For example, by using $L1$ and $L2$ to specify PS1 (in ADAG 1f), the error term would contain only information from the variables that are instruments for A (i.e. $L0$, U_{A1} , and the unspecified unmeasured common cause(s) of A and $L3$ represented by the dashed bidirected arc).

Conclusion

Balancing scores, such as PS and DRS, can be depicted on ADAGs [3, 10]. Due to space constraints, we focused on PS in DAGs, although similar results can be developed for the DRS. Drawing ADAGs should be used after drawing and using the unmodified DAG for identification. In this commentary, we highlighted only some implications of PS-augmented

graphs, but we hope more work will be done on extending them. We argue that ADAGs can be more useful than plain DAGs in empirical causal inference work and deserve further research to help us understand what we can and cannot do with them.

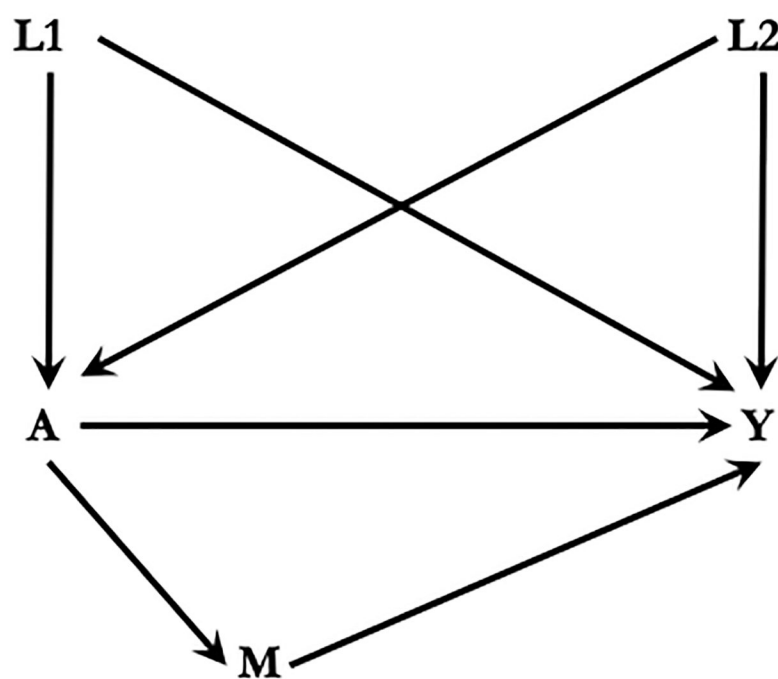
Funding

OAA was supported in part by the Centers for Disease Control and Prevention (CDC) grant agreement number T42OH008412, the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH) under the UCLA Clinical and Translational Science Institute grant number UL1TR001881, and the Karen Toffler Charitable Trust. The National Cancer Institute (NCI/NIH) grant number T32CA009142 supported MMC. Funding does not imply endorsement by or the views of the funders.

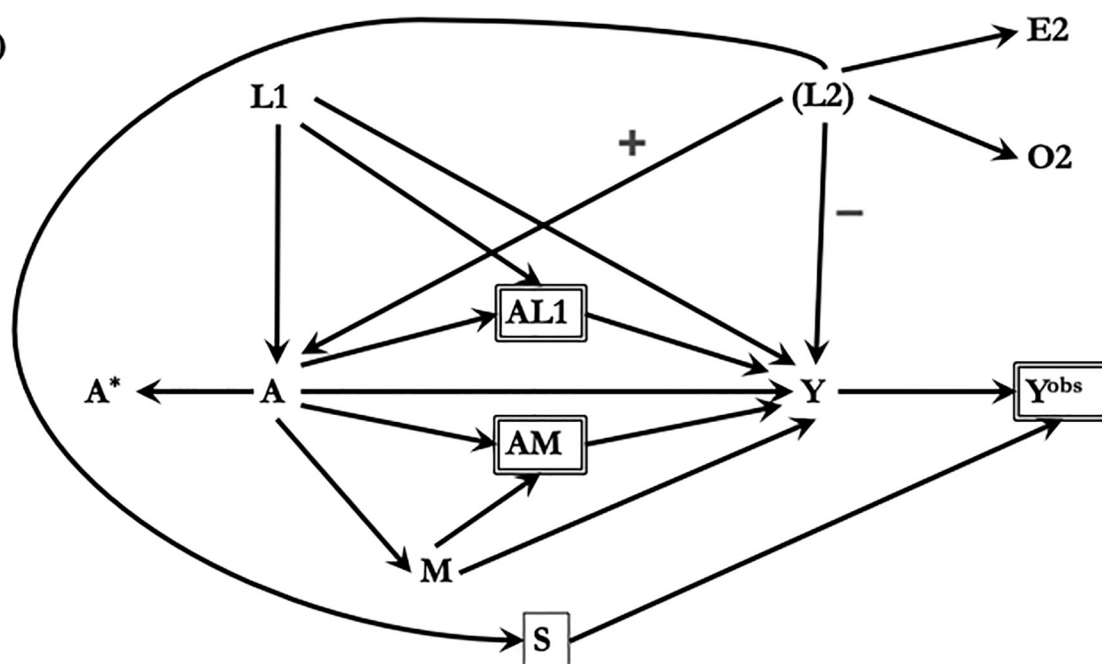
References

1. Geffner H, Dechter, Halpern JY (Editors). Probabilistic and Causal Inference: The Works of Judea Pearl. Association for Computing Machinery, 2022.
2. 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;50:620–632. [PubMed: 33330936]
3. Mansournia M, Nazemipour M, Etminan M. Balancing scores and causal diagrams. *Int J Epidemiol* 2025;54:dyaf114. [PubMed: 40623152]
4. Robins JM. Data, design, and background knowledge in etiologic inference. *Epidemiology* 2001;12:313–320. [PubMed: 11338312]
5. Wardle MT, Reavis KM, Snowden JM. Measurement error and information bias in causal diagrams: mapping epidemiological concepts and graphical structures. *Int J Epidemiol* 2025;53:dyae141.
6. Arah OA. Augmenting causal diagrams with effect modification, interaction, and other parametric information. Abstract #283 in the Abstract Book of the 48th Annual Meeting of the Society for Epidemiologic Research, June 16–19, 2015. <https://github.com/oacarah/causal-diagrams.git> (9 October 2025, date last accessed).
7. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615–25. [PubMed: 15308962]
8. Arah OA. Analyzing selection bias for credible causal inference: When in doubt, DAG it out. *Epidemiology* 2019;30:517–520. [PubMed: 31033691]
9. Berrie L, Arnold KF, Tomova GD, Gilthorpe MS, Tennant PWG. Depicting deterministic variables withing directed acyclic variables: an aid for identifying and interpreting causal effects involving derived variables. *Am J Epidemiol* 2025;194:469–479. [PubMed: 38918044]
10. Arah OA. Using causal diagrams optionally augmented with functional mapping to understand covariate balance in propensity score models. Abstract #445 in the Abstracts of the 42nd Annual Meeting of the Society for Epidemiologic Research, June 23–26, 2009. *Am J Epidemiol* 2009;169(suppl 11):S112. <https://github.com/oacarah/causal-diagrams.git> (9 October 2025, date last accessed).

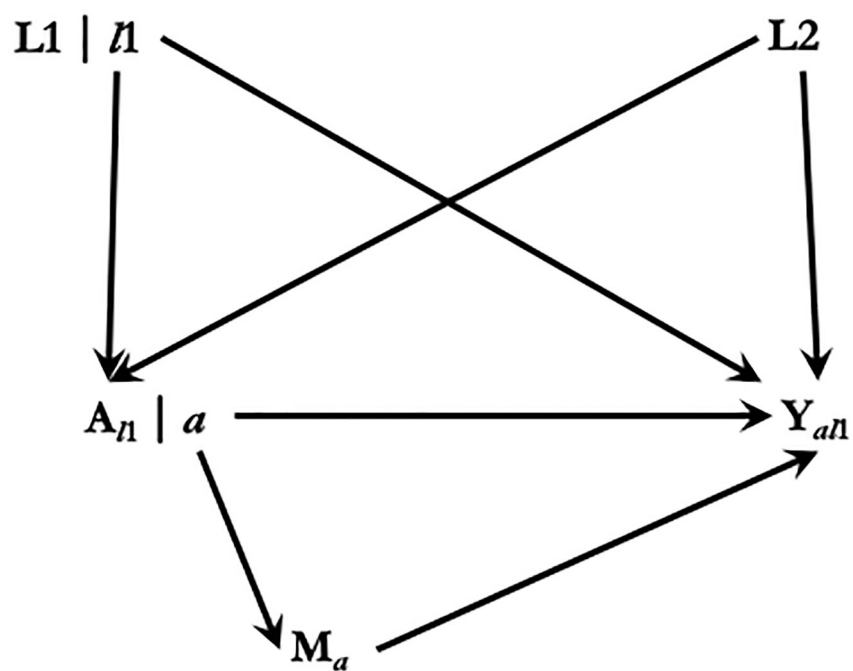
(a)



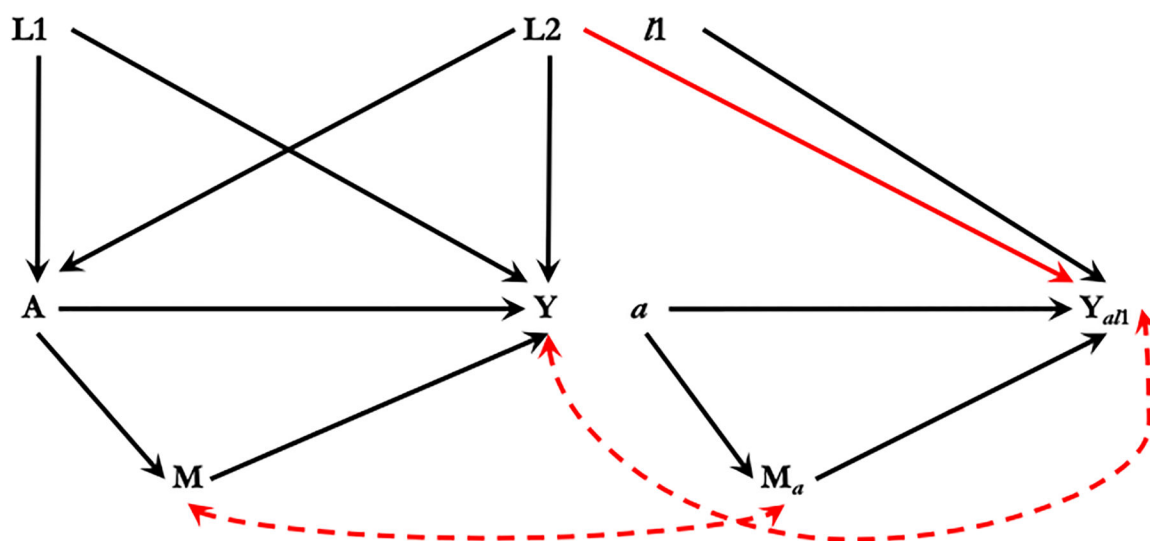
(b)



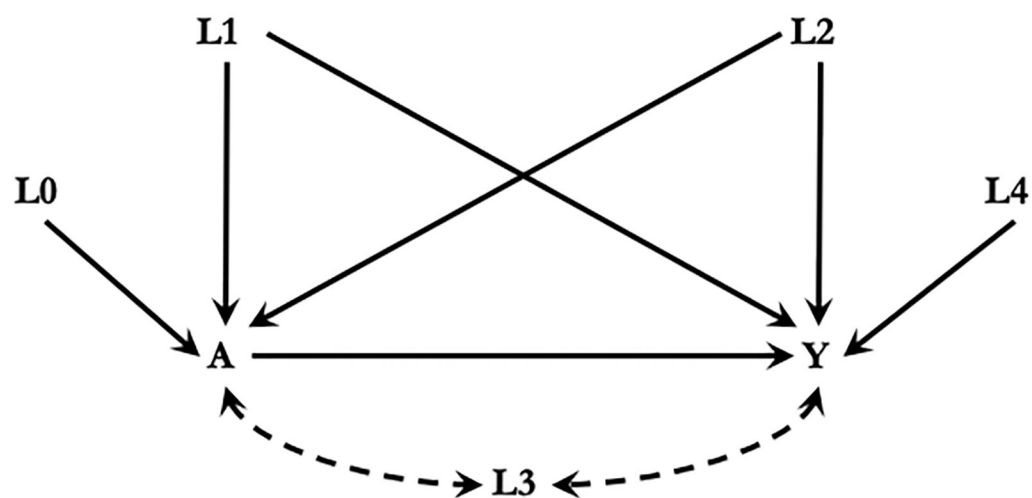
(c)



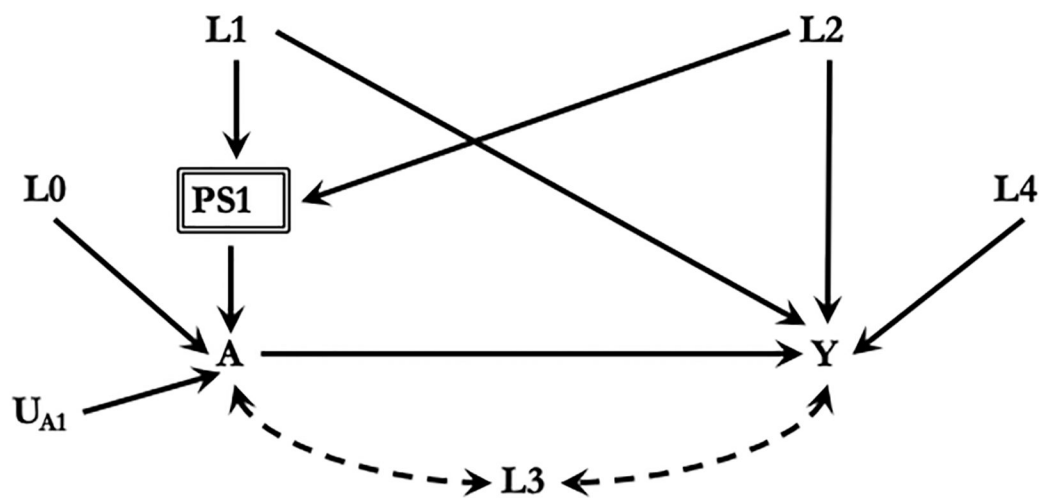
(d)



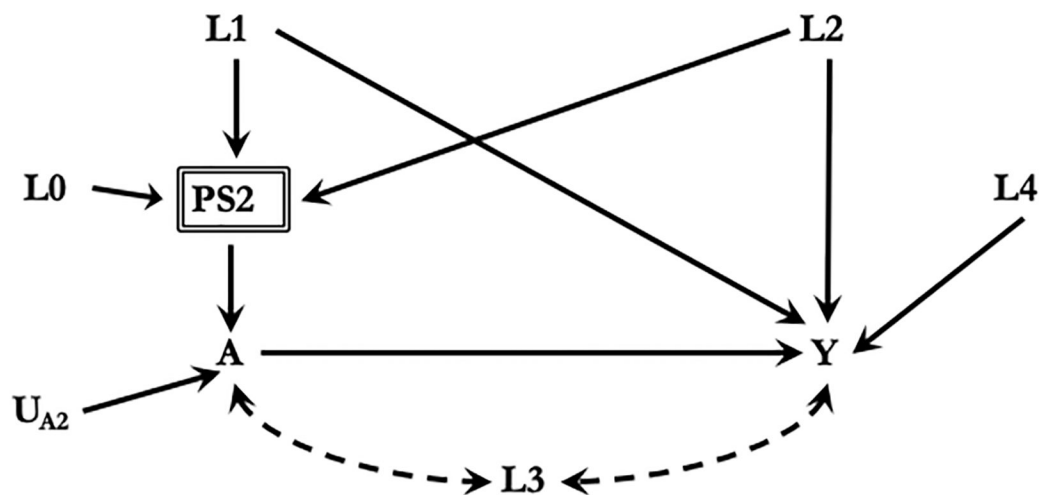
(e)



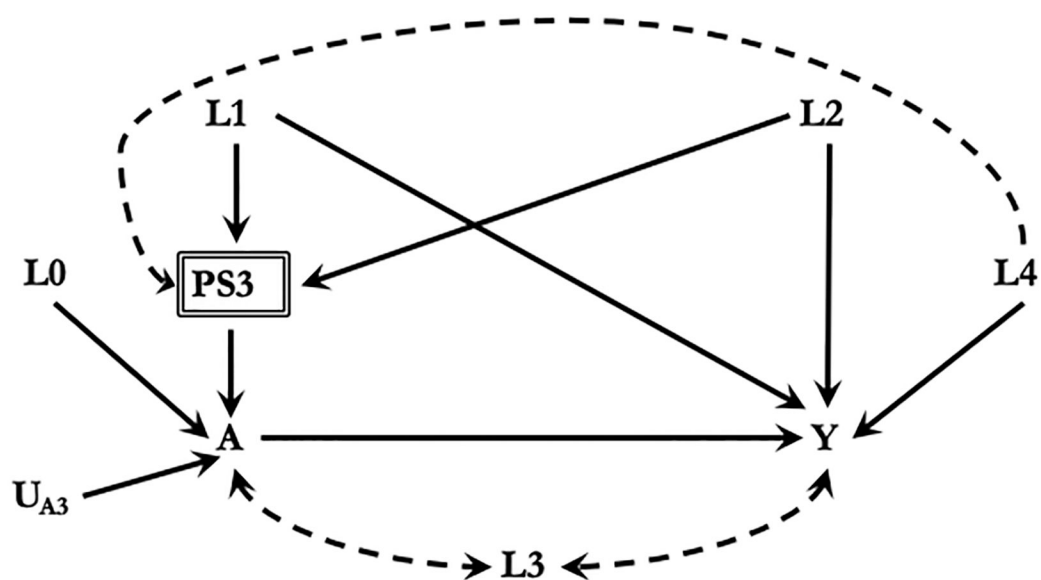
(f)

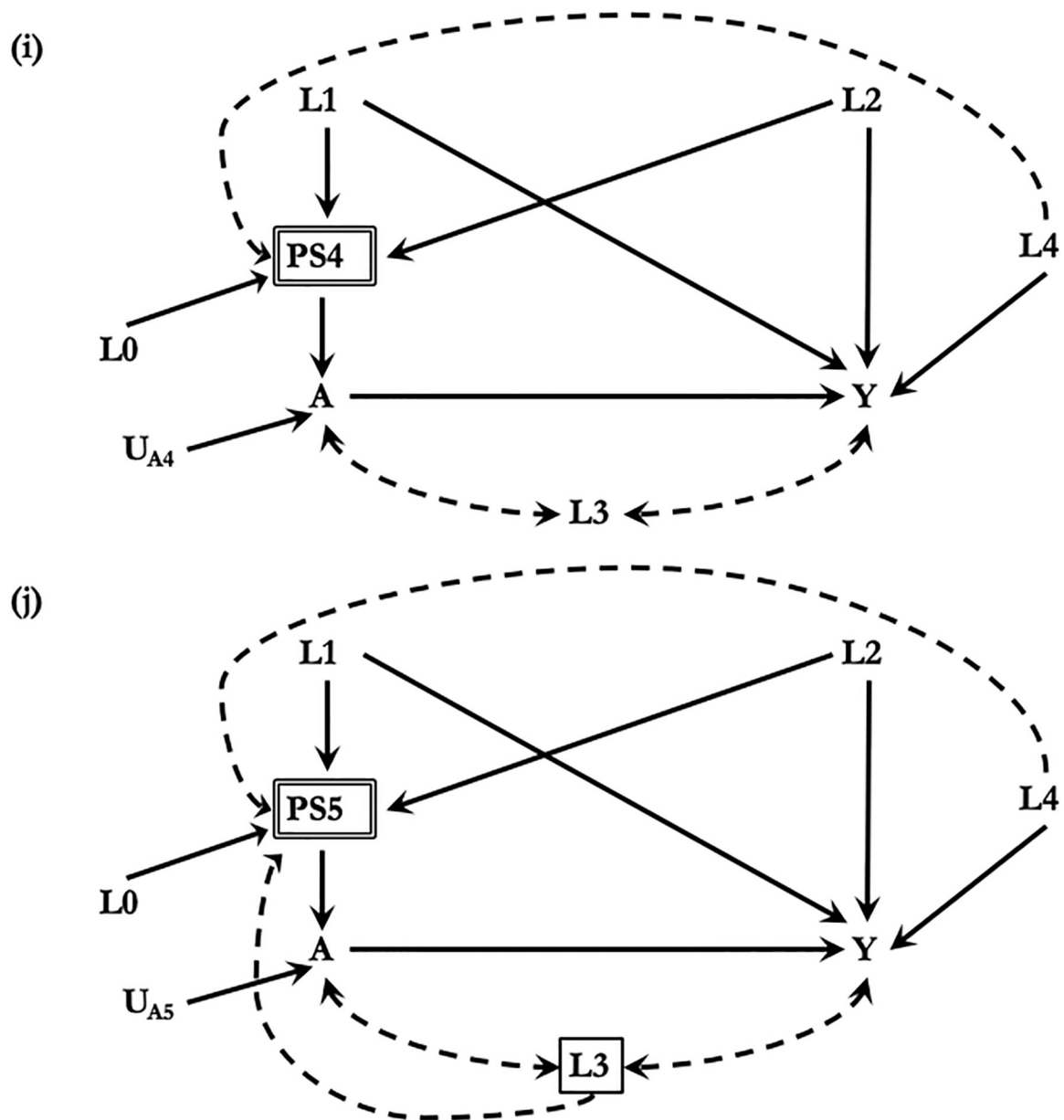


(g)



(h)



**Figure 1:**

Nonaugmented and augmented causal diagrams or directed acyclic graphs

(a) Nonaugmented directed acyclic graph (DAG) for the effect of A on Y . (b) Augmented DAG (ADAG) showing: (i) non-differential independent measurement error in A with node A^* , (ii) effect measure modification by L_1 with AL_1 as a deterministic product node, (iii) interaction between A and M in causing Y with a similar product node augmentation, (iv) unobserved nature of L_2 with parentheses but observable proxies E_2 and O_2 which can be used for proximal causal inference, (v) a positive sign on arrow from unmeasured confounder L_2 to A and a negative sign on the arrow from L_2 to Y indicating the direction of those links and hence of uncontrolled confounding under some monotonicity assumptions, and (vi) selection of observations in which Y is only observed (Y^{obs}) whenever the selection node $S = 1$, and selection is a consequence of an unobserved parent (L_2) of Y .

Furthermore, in ADAG 1b, adding the AM node allows us to see how the total effect of A on Y can be decomposed into four paths or component effects with respect to mediation and interaction: the direct $A \rightarrow Y$ path (controlled direct effect), $A \rightarrow AM \rightarrow Y$ path not through M (reference interaction effect), $A \rightarrow M \rightarrow AM \rightarrow Y$ path (mediated interaction effect), and $A \rightarrow M \rightarrow Y$ path not through AM (pure indirect effect). An additional 5th path $A \rightarrow AL1 \rightarrow Y$ represents a potential $L1$ -modified direct effect of A on Y . These component direct paths, although implied by DAG 1a, are not apparent without augmentation. (c) Corresponding single-world intervention graph (SWIG) with some observed nodes and consequences of joint interventions setting $L1 = \bar{1}$ and $A = a$. (d) Corresponding twin network graph with all observed nodes and consequences of intervention setting $L1 = \bar{1}$ and $A = a$, with shared parents between networks depicted with red arrows. (e) DAG that could be used to study the effect of exposure A on outcome Y , showing parents of A and Y ($L1$ and $L2$), parent of A ($L0$), parent of Y ($L4$), a collider ($L3$), and the unknown parents (U_A) of A being independent of the unknown parents of other variables. (f) ADAG that includes a propensity score PS1 for A specified as a function of $L1$ and $L2$, shown as a deterministic node with a double box. (g) ADAG that includes a propensity score PS2 for A specified using $L0$, $L1$, and $L2$. (h) ADAG that includes a propensity score PS3 for A specified using $L1$, $L2$, and $L4$, with a dashed arrow from $L4$ to PS3 indicating that $L4$ was only used to define PS3 but is not a parent of A . (i) ADAG that denotes a propensity score PS4 for A as a function of $L0$, $L1$, $L2$, and $L4$. (j) ADAG that depicts a propensity score PS5 for A specified as a function of $L1$, $L2$, $L3$, and $L4$, with a directed dashed arrow from $L3$ to PS5 indicating that $L3$ was used to define PS5 but is not an ancestor of A . The resulting data are unfaithful to the ADAG in that the independence of A and $L3$, conditional on PS5, is not evident using d-separation on the ADAG. $L3$ should be avoided in estimating a propensity score because conditioning on PS5, a descendant of collider $L3$, induces a non-causal association between A and Y , which is evident from the ADAG.

Table 1:

Using augmented directed acyclic graphs (ADAGs) to assess variable selection,^a variable balance, and statistical efficiency for different propensity scores used in estimating the effect of A on Y .

Propensity scores with different variable selections	Variable balance before and after applying each propensity score (PS): whether A and each L are d-separated before and after applying the PS <i>b</i> Before applying the PS: $A \perp\!\!\!\perp L$ in DAG 1e After applying the PS: $A \perp\!\!\!\perp L$ in ADAGs 1f to 1j					Propensity score can be used to close open backdoors between A and Y	Can amplify uncontrolled confounding by including instrumental variable(s) in PS	Statistical efficiency ranking (smallest asymptotic variance) ^c
	$L0$ directly causes exposure only	$L1$ causes exposure and outcome	$L2$ causes exposure and outcome	$L3$ does not cause exposure or outcome	$L4$ directly causes outcome only			
PS1: $P(A=1 L1, L2)$ ADAG 1f	Before PS1: No After PS1: No	Before PS1: No After PS1: Yes	Before PS1: No After PS1: Yes	Before PS1: No After PS1: No	Before PS1: Yes After PS1: Yes	Yes	No	2 (second best) Reduces variance: $L1, L2$ Magnifies variance: none
PS2: $P(A=1 L0, L1, L2)$ ADAG 1g	Before PS2: No After PS2: Yes	Before PS2: No After PS2: Yes	Before PS2: No After PS2: Yes	Before PS2: No After PS2: No	Before PS2: Yes After PS2: Yes	Yes	Yes	4 Reduces variance: $L1, L2$ Magnifies variance: $L0$
PS3: $P(A=1 L1, L2, L4)$ ADAG 1h	Before PS3: No After PS3: No	Before PS3: No After PS3: Yes	Before PS3: No After PS3: Yes	Before PS3: No After PS3: No	Before PS3: Yes After PS3: Yes	Yes	No	1 (best) Reduces variance: $L1, L2, L4$ Magnifies variance: none
PS4: $P(A=1 L0, L1, L2, L4)$ ADAG 1i	Before PS4: No After PS4: Yes	Before PS4: No After PS4: Yes	Before PS4: No After PS4: Yes	Before PS4: No After PS4: No	Before PS4: Yes After PS4: Yes	Yes	Yes	3 Reduces variance: $L1, L2, L4$ Magnifies variance: $L0$
PS5: $P(A=1 L0, L1, L2, L3, L4)$ ADAG 1j	Before PS5: No After PS5: Yes	Before PS5: No After PS5: Yes	Before PS5: No After PS5: Yes	Before PS5: No After PS5: Yes	Before PS5: Yes After PS5: Yes	Opens a collider-biasing path between A and Y by conditioning on PS5, a descendent of collider $L3$	Yes	5 (worst) Reduces variance: $L1, L2$ Magnifies variance: $L0, L3$

^aVariable selection using the backdoor criterion in DAG 1e: The minimum set of deconfounders is $\{L1, L2\}$, which can close the open backdoors $A \leftarrow L1 \rightarrow Y$ and $A \leftarrow L2 \rightarrow Y$.

^bThe symbol $\perp$ is used to indicate 'independent of'.

^cAn optimal or efficient valid adjustment set (e.g., $L1, L2$, and $L4$ in DAG 1e) is the set of the parents of the outcome Y , excluding the consequences of exposure A that cause Y . This set contains the so-called 'precision variables' (such as set $L4$) and yields the smallest asymptotic variance relative to the other variable adjustment sets in many standard parametric and non-parametric models [see, for

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

example, Witte J, Henckel L, Maathuis MH, Didelez V. On efficient adjustment in causal graphs. *J Machine Learning Res* 2020;21:1–45]. Including a parent (e.g., $L0$ in DAG 1e) of exposure A that is d-separated by A from Y worsens efficiency relative to the efficient valid set. Therefore, a PS model that includes the most parents of A only) would be expected to yield the smallest variance of the estimated effect of A on Y using a PS method. A PS-adjusted model, however, would yield a larger variance compared to using an outcome model that directly adjusts for $L1$, $L2$, and $L4$ (i.e., without a PS). This occurs because, when confounding equivalent sets exist for a DAG (e.g., PS3 and $\{L1, L2, L4\}$ for ADAG 1h), adjusting for the variable set ($L1, L2, L4$) closest to the outcome Y tends to result in better statistical efficiency. Adjusting for variables closest (or those that are closer) to the exposure variable A tends to worsen statistical efficiency.