Causal diagrams, expert opinion and structure learning: vetting the expert

CJ Oates[*1], **J Kasza**[*2,3], JA Simpson[3,4], AB Forbes[2,3]
* Equal joint first authors

jessica.kasza@monash.edu

1. University of Technology Sydney
2. Monash University
3. Victorian Centre for Biostatistics (ViCBiostat)
4. University of Melbourne

Estimating effect of an exposure on an outcome

- **Use data to investigate the relationship between an exposure and an outcome.**

  e.g. What is the relationship between smoking and adult asthma?

- How can we **best select the set of confounders** to adjust for in our estimation of the exposure-outcome relationship?

Simple approach:

- Ask an expert to list confounders, adjust for these in outcome regression model or propensity score model.

A better approach:

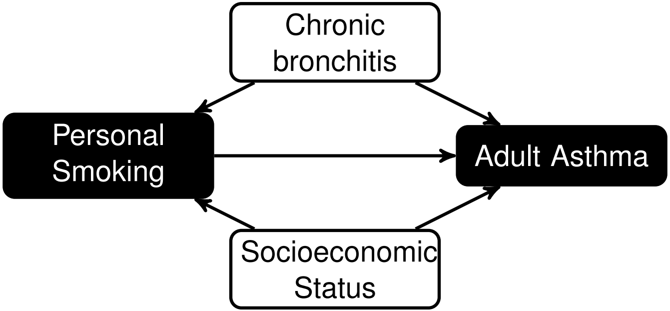
- Get the expert to draw a **causal diagram...**

Causal diagrams and DAGs

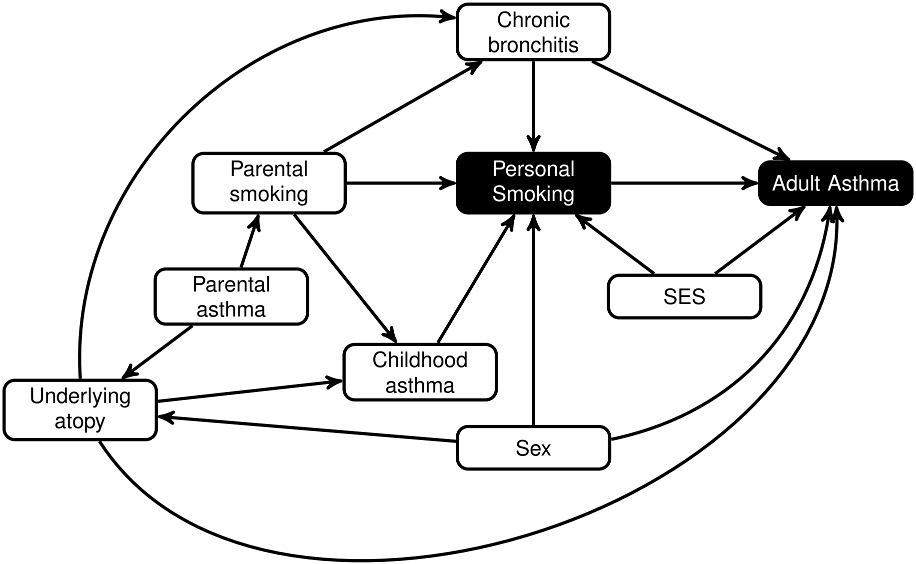
Causal diagrams display relationships using directed acyclic graphs (DAGs)

$$\mathcal{G} = (V, E)$$

- $V =$ set of variables (nodes);
- $E =$ set of directed edges between nodes (indicating direct causes);
- missing edges indicate the absence of direct causal relationships.



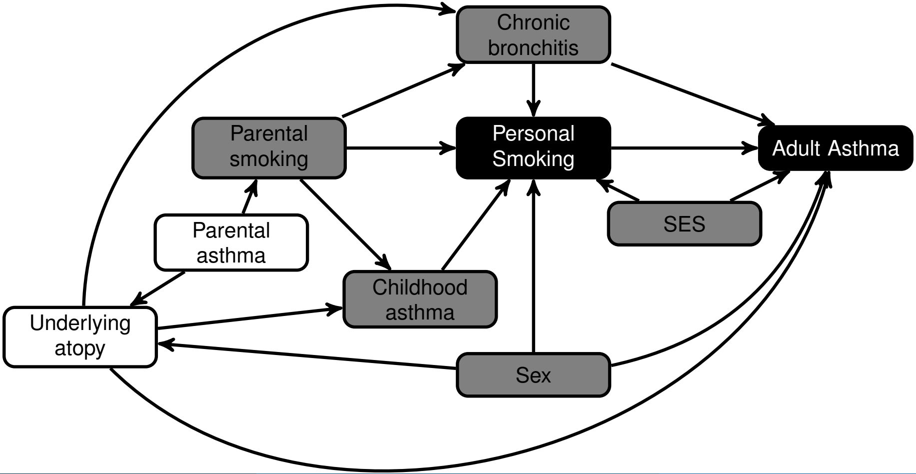
A more realistic DAG: from Williamson et al. *Respirology*, 2014.



Why is a DAG a better approach?

Can help to prevent bias from over- or under-adjustment; increase efficiency.

- Apply graph-theoretic rules to determine adjustment set e.g. `dagitty.net`.



What if the expert got the DAG wrong?

True DAG: $\mathcal{G}_{\text{true}} = (V_{\text{true}}, E_{\text{true}})$ **Expert DAG:** $\mathcal{G}_{\text{expert}} = (V_{\text{expert}}, E_{\text{expert}})$

- We assume the set of variables is correctly specified, but the edge set may be misspecified.
  - $V_{\text{true}} = V_{\text{expert}}$
  - If $E_{\text{true}} \subseteq E_{\text{expert}}$, $\mathcal{G}_{\text{expert}}$ is valid for causal inference.

But how can we tell if the expert's DAG is valid?

Structure learning and the expert

Structure learning algorithm (unconstrained):

Input: n samples of $(X_1, X_2, \ldots, X_p)$

Output: DAG[1] on $(X_1, X_2, \ldots, X_p)$

Apply structure learning algorithms to the data, compare result to the expert's DAG[2]

[1] Not quite true: it will actually find an **equivalence class** of DAGs...

[2] For example: Meek, *Causal inference and causal explanation with background knowledge*. UAI 1995

Structure learning and the expert

Structure learning algorithm (unconstrained):

Input: n samples of $(X_1, X_2, \ldots, X_p)$

Output: DAG[1] on $(X_1, X_2, \ldots, X_p)$

Apply structure learning algorithms to the data, compare result to the expert's DAG[2]

PROBLEMS!

- These algorithms are quite unstable: small perturbations of the data may lead to very different structures.
- As $n \to \infty$, these algorithms will find the true underlying DAG[1], **BUT** behaviour for realistic sample sizes can be poor...
  - Large space, low statistical power to detect associations.

[1] Not quite true: it will actually find an **equivalence class** of DAGs...

[2] For example: Meek, *Causal inference and causal explanation with background knowledge*. UAI 1995

Our proposal: **vetting** (*Constrained* structure learning)

Vetting= **V**alidation of **E**xpert **T**opology

Input: n samples of $(X_1, X_2, \ldots, X_p)$ **plus the expert's DAG**

Output: DAG on $(X_1, X_2, \ldots, X_p)$: an extended version of the expert's DAG

- Constrain structure learning algorithms so that edges specified by the expert are always included.
  - Only consider super-graphs of the expert's graph: additional edges necessary?
  - Super-graphs are valid for causal inference.
- Requires development of the theory of *vetting equivalence classes*...

***How* can the expert go wrong?**

The good, the bad, and the ugly: types of expert errors

True DAG:

$$X \longrightarrow Y \longrightarrow Z$$

The good, the bad, and the ugly: types of expert errors

True DAG:

$$X \longrightarrow Y \longrightarrow Z$$

1. “Essentially correct”: truth is contained in the expert’s graph

$$X \longrightarrow Y \longrightarrow Z$$

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend left] & Y \arrow[r] & Z \end{tikzcd}$$

The good, the bad, and the ugly: types of expert errors

True DAG:

$$X \longrightarrow Y \longrightarrow Z$$

1. "Essentially correct": truth is contained in the expert's graph

$$X \longrightarrow Y \longrightarrow Z \qquad \begin{tikzcd} X \arrow[r] \arrow[rr, bend left] & Y \arrow[r] & Z \end{tikzcd}$$

2. "Weakly incorrect": a super-graph of the expert's graph that contains the truth

$$X \longrightarrow Y \qquad Z \qquad \begin{tikzcd} X \arrow[rr, bend left] & Y \arrow[r] & Z \end{tikzcd}$$

The good, the bad, and the ugly: types of expert errors

True DAG:

$$\begin{tikzcd} X \arrow[r] & Y \arrow[r] & Z \end{tikzcd}$$

1. "Essentially correct": truth is contained in the expert's graph

$$\begin{tikzcd} X \arrow[r] & Y \arrow[r] & Z \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r] \arrow[rr, bend left] & Y \arrow[r] & Z \end{tikzcd}$$

2. "Weakly incorrect": a super-graph of the expert's graph that contains the truth

$$\begin{tikzcd} X \arrow[r] & Y & Z \end{tikzcd} \qquad \begin{tikzcd} X \arrow[rr, bend left] & Y \arrow[r] & Z \end{tikzcd}$$

3. "Strongly incorrect": no extension of the expert's graph contains the truth

$$\begin{tikzcd} X & Y \arrow[l] \arrow[r] & Z \end{tikzcd} \qquad \begin{tikzcd} X & Y & Z \arrow[ll, bend right] \end{tikzcd}$$

The good, the bad, and the ugly: types of expert errors

True DAG:

$$X \longrightarrow Y \longrightarrow Z$$

1. **"Essentially correct"**: truth is contained in the expert's graph

$$X \longrightarrow Y \longrightarrow Z \qquad\qquad \begin{tikzcd} X \arrow[r] \arrow[rr, bend left] & Y \arrow[r] & Z \end{tikzcd}$$

2. **"Weakly incorrect"**: a super-graph of the expert's graph that contains the truth

$$X \longrightarrow Y \qquad Z \qquad\qquad \begin{tikzcd} X \arrow[rr, bend left] & Y \arrow[r] & Z \end{tikzcd}$$

3. "Strongly incorrect": no extension of the expert's graph contains the truth

$$X \longleftarrow Y \longrightarrow Z \qquad\qquad \begin{tikzcd} X & Y & Z \arrow[ll, bend right] \end{tikzcd}$$

Simulation study: Vetting the Asthma DAG

Generate binary data from the Asthma DAG: $\mathcal{G}_{true}$

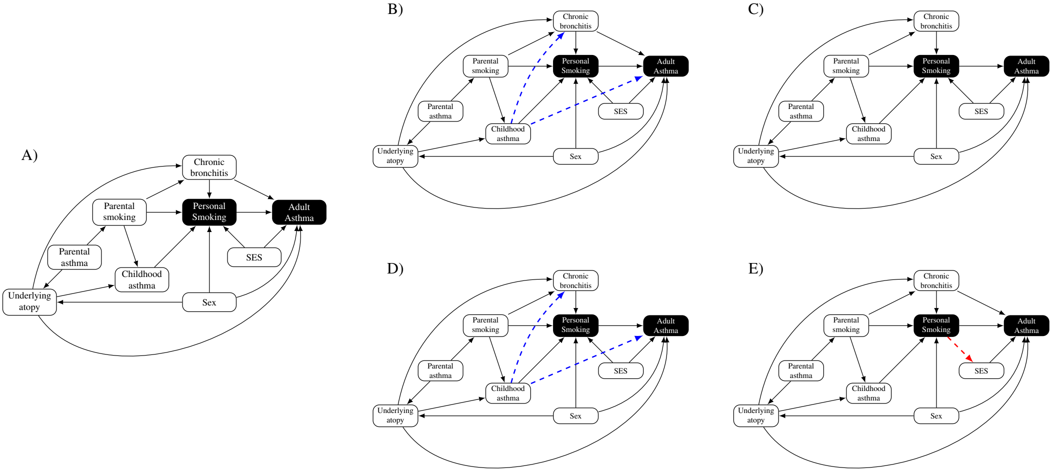
- 50 data sets for each sample size $n = \{10, 50, 100, 250, 10000\}$.

1. Apply vetting to an 'expert-elicited' DAG
2. Apply unconstrained structure learning

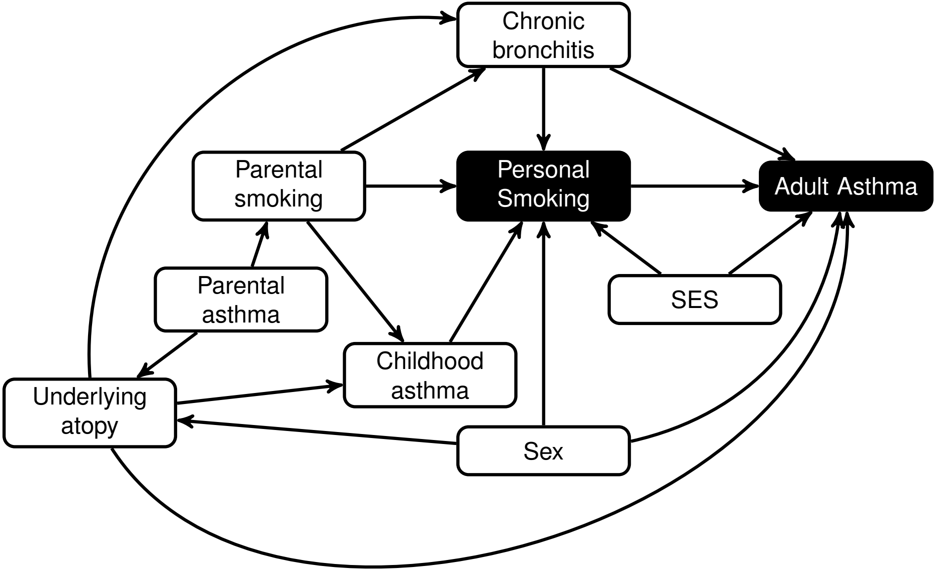
For each learned equivalence class of DAGs:

1. Is the returned DAG equivalence class *correct*?
   - Calculate $P(\mathcal{G}_{\text{true}} \subseteq \mathcal{G}_{\text{learned}})$: super-model of true causal DAG valid for causal inference.
2. Is the estimate of the average causal effect of personal smoking on adult asthma *unbiased*?
   - Use inverse probability of treatment weighting to estimate the effect
   - Calculate squared error of estimate.

The true and expert DAGs

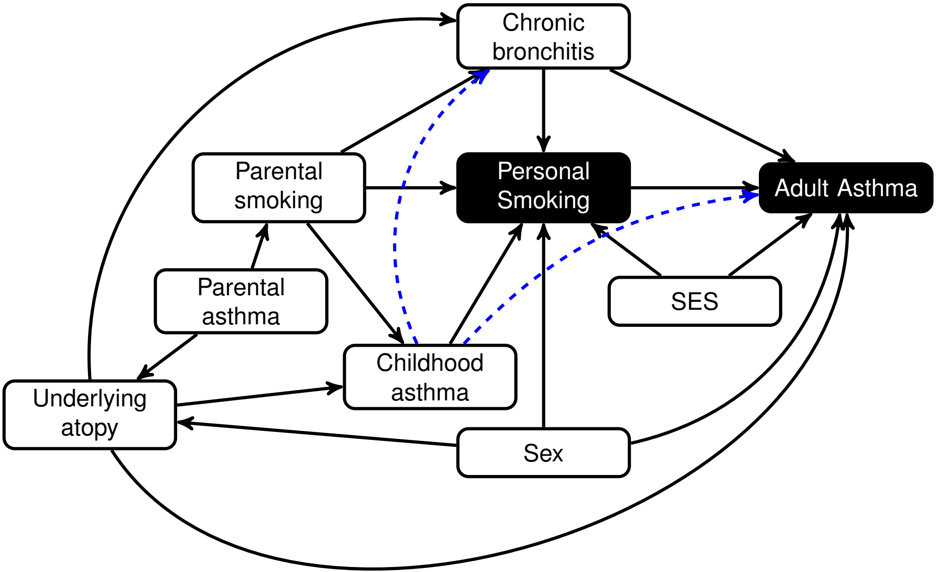


A) True DAG

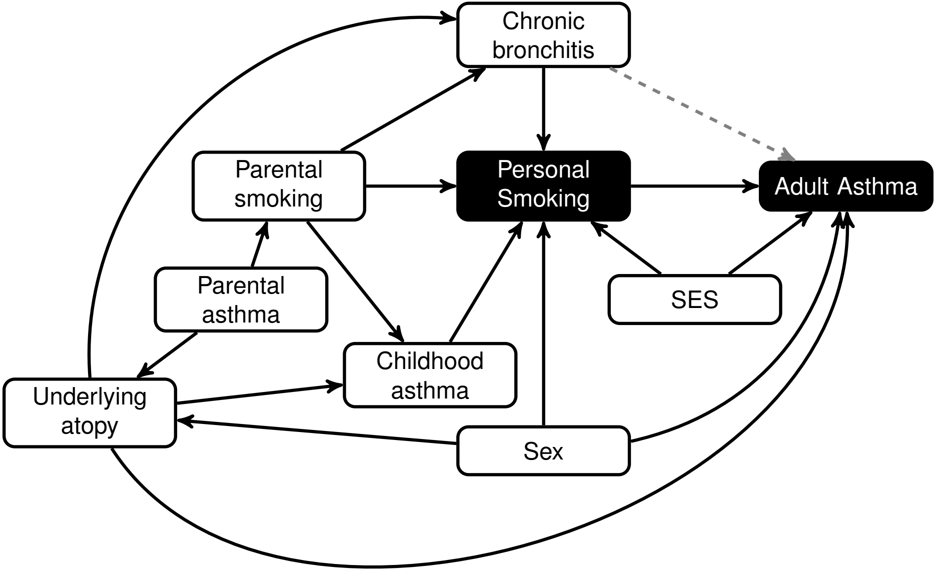


Chronic bronchitis
Parental smoking
Personal Smoking
Adult Asthma
Parental asthma
SES
Childhood asthma
Underlying atopy
Sex

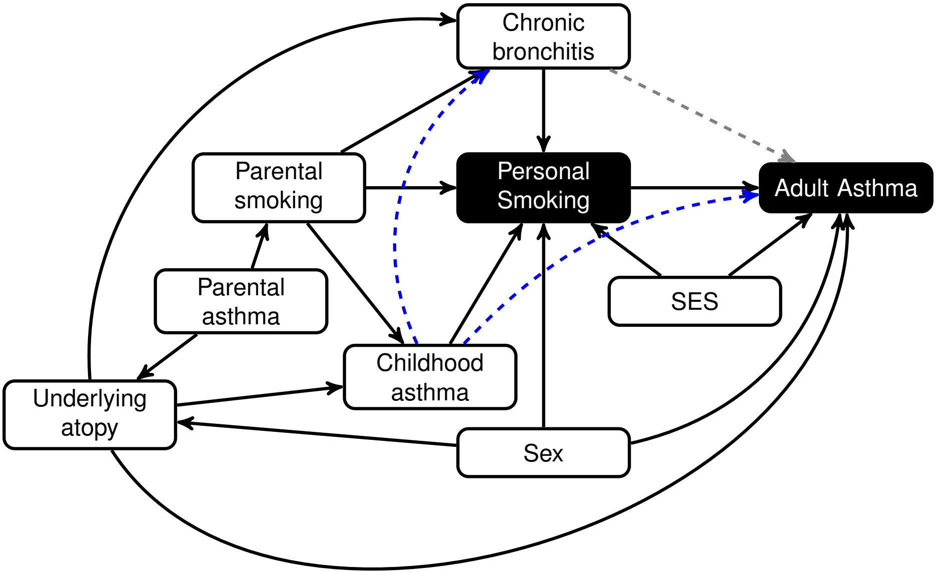
B) Expert essentially correct (Good!)



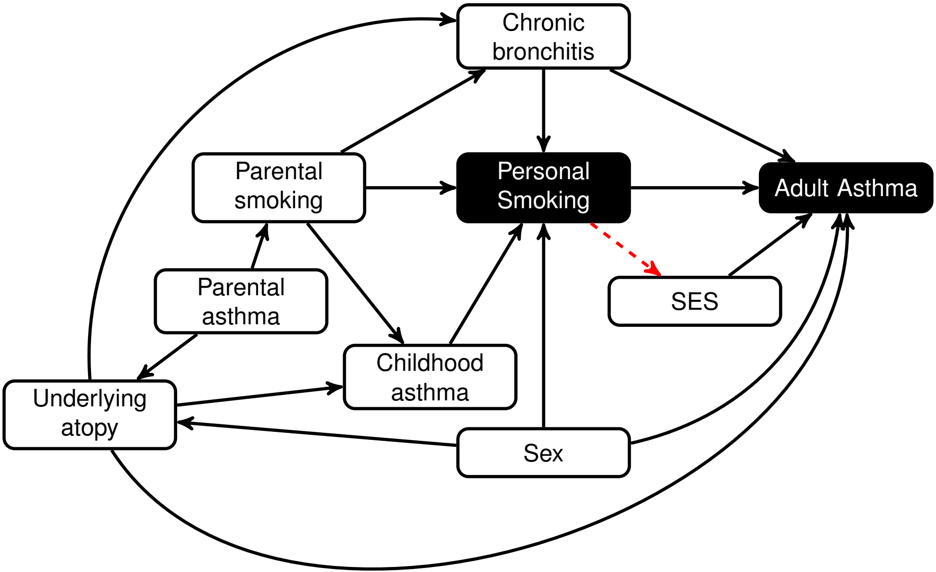
C) Expert weakly incorrect I (Ugly!)



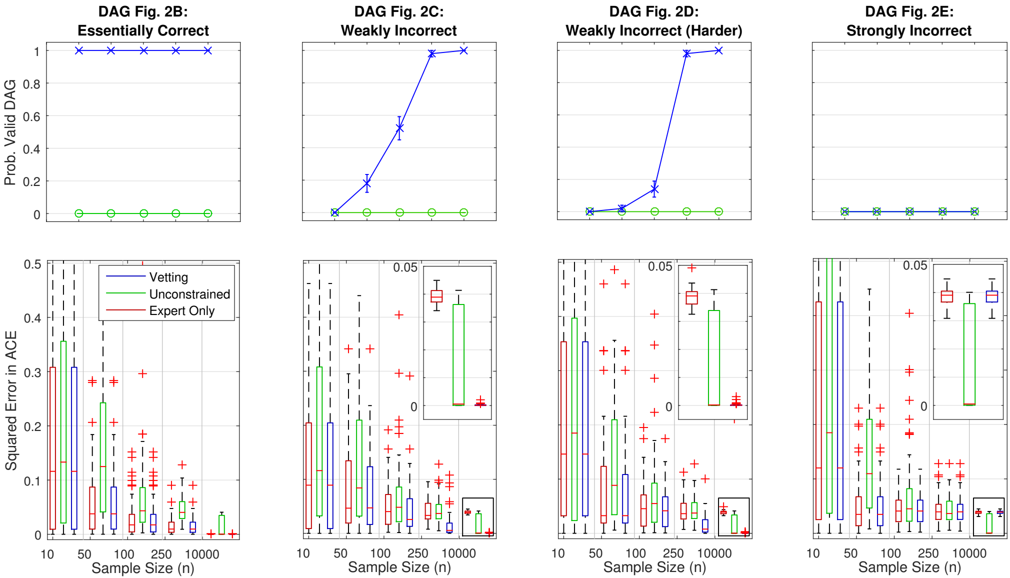
D) Expert weakly incorrect II (Ugly!)



E) Expert strongly incorrect (Bad!)



Simulation study results

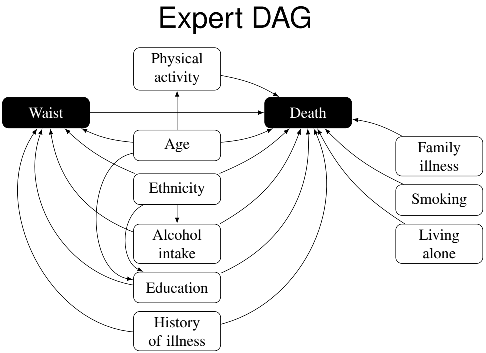


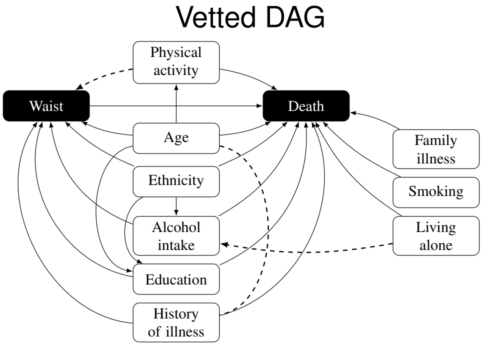
Real data example: Melbourne Collaborative Cohort Study

What is the total causal effect of waist circumference on mortality?

- Random subset of 9000 male participants from MCCS
- Apply vetting to the expert's DAG, using a randomly selected sub-subset of size $n = 1000$.
- Use $n = 9000$ data to estimate total causal effect of waist circumference on mortality using IPTW, adjusting for variables as indicated by the $\{\text{expert}, \text{vetted}\}$ DAG.

Estimating the effect of waist circumference $\geq$ 102cm on death





Adjustment	Odds Ratio (95% CI)
No adjustment	2.08 (1.79, 2.43)
Expert DAG	1.53 (1.31, 1.80)
Vetted DAG	1.51 (1.29, 1.78)

Vetting: take-home messages

Causal diagrams, using the language of DAGs, are a useful tool for adjustment set selection.

- Expert opinion is invaluable in the construction of a DAG. But...
  - DAG construction is difficult!
  - Erroneous DAGs can lead to invalid inference.

Vetting: augmentation of the expert's DAG using structure learning.

- Assumptions: all necessary variables measured and included and a supergraph of the expert's DAG contains the true DAG.
- Automated procedures balanced with expert knowledge.
- Robustness against certain types of expert errors.

References

- Meek C, Causal inference and causal explanation with background knowledge. *UAI 1995*.
- Oates CJ, Kasza J, Simpson JA, Forbes A. Repair of partly misspecified causal diagrams. *Epidemiology*. Accepted: to appear 2017.
- Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology*. 2008;8:70.
- Williamson EJ, Aitken Z, Lawrie J, et al. Introduction to causal diagrams for confounder selection. *Respirology* 2014;19(3):303-11.


SAVE THE DATE
Joint International Society for Clinical Biostatistics and Australian Statistical Conference 26-30 August 2018
ISCB
ASC18
26-30 AUGUST 2018
MELBOURNE, AUSTRALIA

Markov & Vetting equivalence classes of DAGs

Not all DAGs encode different sets of conditional independence relationships...
For example,

$$\{X \rightarrow Y \rightarrow Z, X \leftarrow Y \leftarrow Z, X \leftarrow Y \rightarrow Z\}$$

all encode that X and Z are conditionally independent given Y.

- This set of DAGs forms a **Markov equivalence class** of DAGs

However, if the expert specifies an edge $X \rightarrow Y$, then

$$\{X \rightarrow Y \rightarrow Z\} \text{ and } \{X \leftarrow Y \leftarrow Z, X \leftarrow Y \rightarrow Z\}$$

form separate vetting equivalence classes.

- Can distinguish between Markov equivalent DAGs using expert information!
- Vetting equivalence is a finer notion than Markov equivalence.

Learning graphical structure: the PC algorithm

Input: Data n samples of $(X_1, X_2, \ldots, X_p)$

Output: DAG on $(X_1, X_2, \ldots, X_p)$

PC algorithm:

Stage 1: Start with a complete undirected graph on $(X_1, X_2, \ldots, X_p)$

- Test for conditional independence: remove edges

Stage 2: Direct edges of the undirected graph using information about conditional independence.

Vetting version of the PC algorithm:

Stage 1: Start with a complete undirected graph on $(X_1, X_2, \ldots, X_p)$

- Test for conditional independence: remove edges, but only consider removal of edges NOT in expert's DAG.

Stage 2: Direct edges of the undirected graph using the expert's DAG and information about conditional independence.

MCCS example: Unconstrained structure learning