A Short Introduction to Directed Acyclic Graphs (DAGs)

Joshua P. Entrop

Läkemedelsverket, Enhet för Effekt och Säkerhet, October 31st, 2024

Karolinska Institutet

Background - Counterfactuals

Counterfactuals

Y_i^x denotes the potential outcome Y of individual i under treatment x.

E.g. my (i) sleep quality tonight (Y) if I ate pasta ($x = 1$) instead of oats ($x = 0$) for dinner.

Problem

- We are interest in $\mathbb{E}[Y^{x=0}] - \mathbb{E}[Y^{x=1}]$
- We know $\mathbb{E}[Y|x = 0] - \mathbb{E}[Y|x = 1]$

Hence, we need a method for translating our contrafactual outcomes of interest into observable quantities.

Solution: Pearl's *do*-calculus, which requires 3 assumptions.

Background - Identifiability Assumption

1. Consistency ($Y_i^x = Y_i$, if $X_i = x$)

The counterfactual outcome Y_i^x corresponds to the observed outcome Y_i if individual i received treatment x in the real world.

Background - Identifiability Assumption

1. Consistency ($Y_i^x = Y_i$, if $X_i = x$)

2. Conditional Exchangeability ($Y^x \perp\!\!\!\perp X | L$)

The counterfactual outcome is independent of the observed treatment given some adjustment set L.

Background - Identifiability Assumption

1. Consistency ($Y_i^x = Y_i$, if $X_i = x$)
2. Conditional Exchangeability ($Y^x \perp\!\!\!\perp X|L$)

3. Positivity ($\mathbb{P}[X = x, L = l] > 0$)

It should in theory be possible to identify both treated and untreated individuals for each possible combination of the variables included in the adjustment set L.

Background - Identifiability Assumption

1. Consistency ($Y_i^x = Y_i$, if $X_i = x$)
2. Conditional Exchangeability ($Y^x \perp\!\!\!\perp X | L$)
3. Positivity ($\mathbb{P}[X = x, L = l] > 0$)

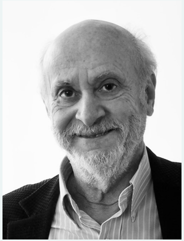
Idetifiability Assumptions and DAGs

1. Consistency ($Y_i^x = Y_i$, if $X_i = x$)
2. **Conditional Exchangeability** ($Y^x \perp\!\!\!\perp X | L$)
3. Positivity ($\mathbb{P}[X = x, L = l] > 0$)

How to asses exchangeability?

Judea Pearl

vs.



Donald B. Rubin

The Basic Ingridiens of DAGs

Mathematically speaking, a DAG is a visual representation of a joint distribution of variables defined by:

- **Nodes**: Variables in our causal network
- **Arrows**: Direction of causation
- Note that the absence of an arrow is a stronger assumption than the presence of it, i.e., complete independence.

The Basic Ingridiens of DAGs

Mathematically speaking, a DAG is a visual representation of a joint distribution of variables defined by:

- **Nodes**: Variables in our causal network
- **Arrows**: Direction of causation
- Note that the absence of an arrow is a stronger assumption than the presence of it, i.e., complete independence.

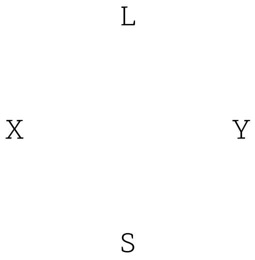


Figure 1. Some nodes

The Basic Ingridiens of DAGs

Mathematically speaking, a DAG is a visual representation of a joint distribution of variables defined by:

- **Nodes**: Variables in our causal network
- **Arrows**: Direction of causation
- Note that the absence of an arrow is a stronger assumption than the presence of it, i.e., complete independence.

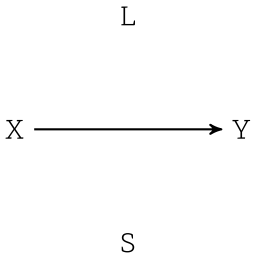


Figure 1. A direct effect

The Basic Ingridiens of DAGs

Mathematically speaking, a DAG is a visual representation of a joint distribution of variables defined by:

- **Nodes**: Variables in our causal network
- **Arrows**: Direction of causation
- Note that the absence of an arrow is a stronger assumption than the presence of it, i.e., complete independence.

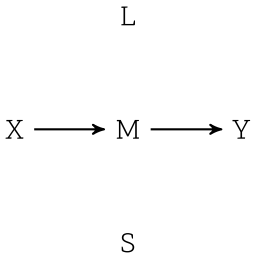


Figure 1. An indirect effect

The Basic Ingridiens of DAGs

Mathematically speaking, a DAG is a visual representation of a joint distribution of variables defined by:

- **Nodes**: Variables in our causal network
- **Arrows**: Direction of causation
- Note that the absence of an arrow is a stronger assumption than the presence of it, i.e., complete independence.

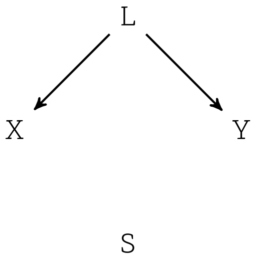


Figure 1. A confounder

The Basic Ingridiens of DAGs

Mathematically speaking, a DAG is a visual representation of a joint distribution of variables defined by:

- **Nodes**: Variables in our causal network
- **Arrows**: Direction of causation
- Note that the absence of an arrow is a stronger assumption than the presence of it, i.e., complete independence.

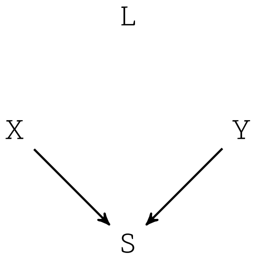


Figure 1. A collider

The Basic Ingridiens of DAGs

Mathematically speaking, a DAG is a visual representation of a joint distribution of variables defined by:

- **Nodes**: Variables in our causal network
- **Arrows**: Direction of causation
- Note that the absence of an arrow is a stronger assumption than the presence of it, i.e., complete independence.

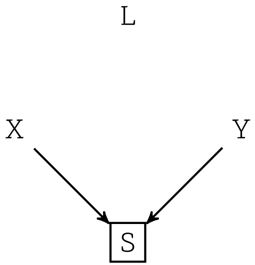


Figure 1. Conditioned on a collider

From Counterfactuals to Reality via *do*-calculus

Back-door criteria

We know as a results from Pearl's *do*-calculus that exchangability holds if there is no open path between X and Y in a DAG in which all outgoing arrows from X are removed.

From Counterfactuals to Reality via *do*-calculus

Back-door criteria

We know as a results from Pearl's *do*-calculus that exchangability holds if there is no open path between X and Y in a DAG in which all outgoing arrows from X are removed.

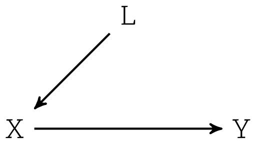


Figure 2. The closed back door

From Counterfactuals to Reality via *do*-calculus

Back-door criteria

We know as a results from Pearl's *do*-calculus that exchangability holds if there is no open path between X and Y in a DAG in which all outgoing arrows from X are removed.

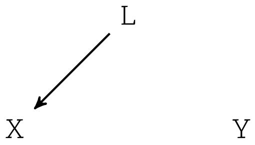


Figure 2. The closed back door

From Counterfactuals to Reality via *do*-calculus

Back-door criteria

We know as a results from Pearl's *do*-calculus that exchangability holds if there is no open path between X and Y in a DAG in which all outgoing arrows from X are removed.

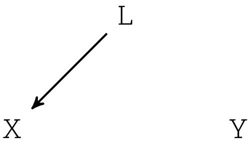


Figure 2. The closed back door

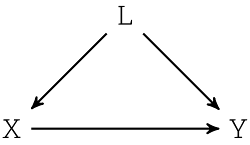


Figure 3. The open back door

From Counterfactuals to Reality via *do*-calculus

Back-door criteria

We know as a results from Pearl's *do*-calculus that exchangability holds if there is no open path between X and Y in a DAG in which all outgoing arrows from X are removed.

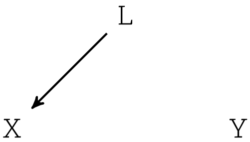


Figure 2. The closed back door

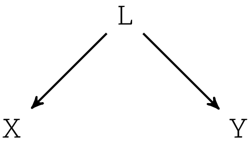


Figure 3. The open back door

From Counterfactuals to Reality via *do*-calculus

Back-door criteria

We know as a results from Pearl's *do*-calculus that exchangability holds if there is no open path between X and Y in a DAG in which all outgoing arrows from X are removed.

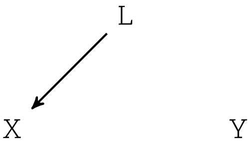


Figure 2. The closed back door

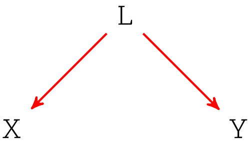


Figure 3. The open back door

A Randomised Experiment

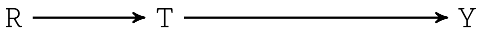


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

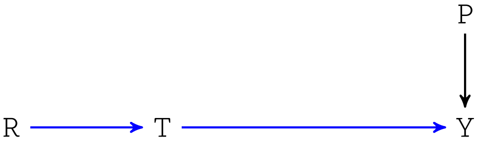


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

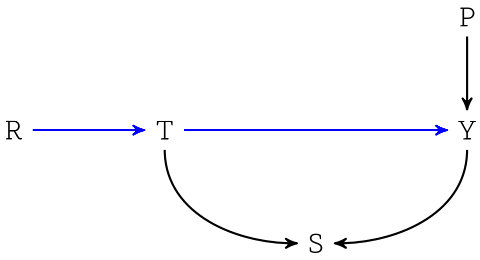


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

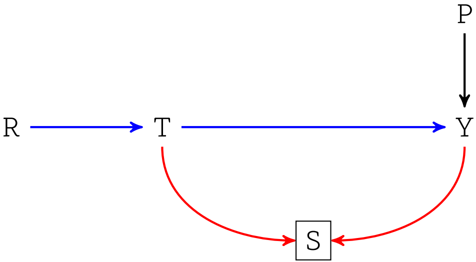


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

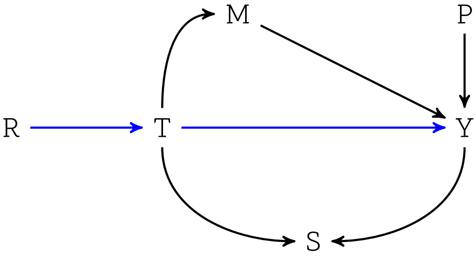


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

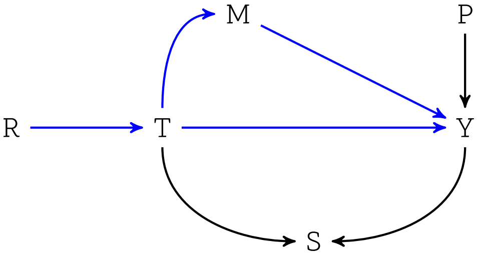


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

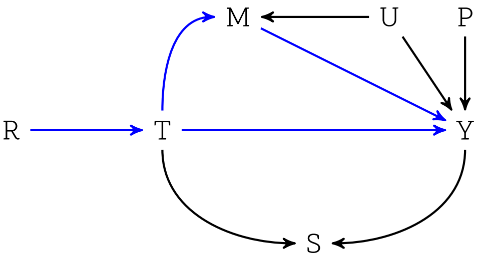


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

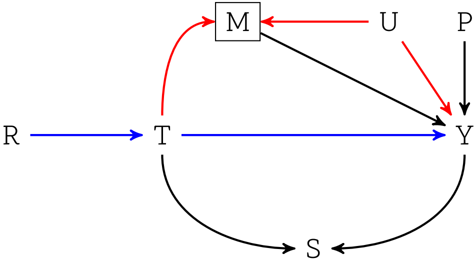


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

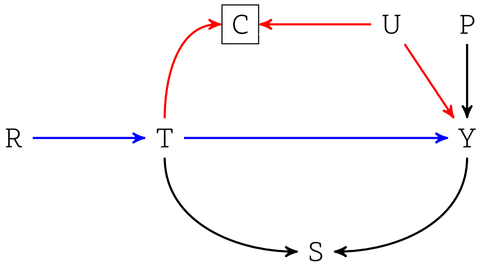


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Randomised Experiment

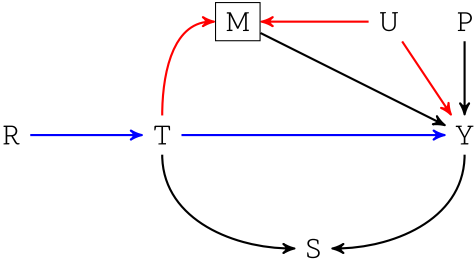


Figure 4. DAG of a randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Blinded Randomised Experiment

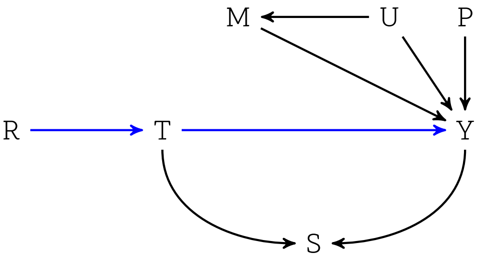


Figure 4. DAG of a blinded randomised experiment

- R: Randomisation
- T: Treatment
- Y: Outcome (weight loss)
- P: Prognostic factors (baseline BMI)
- S: Collider (fatigue)
- M: Mediator (physical activity)
- U: Confounder (SES)
- C: Drop out

A Blinded Randomised Experiment with ICEs

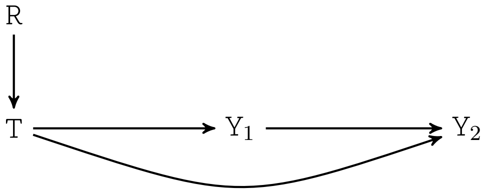


Figure 5. DAG of an ICE mechanism

- R: Randomisation
- T: Treatment
- C: ICE (treatment discontinuation)
- Y_t: Outcome at time t
- U: Confounder (behavioural trait)
- M_{y2}: A mechanism for assigning a value to Y.

A Blinded Randomised Experiment with ICEs

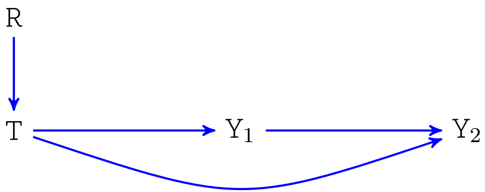


Figure 5. DAG of an ICE mechanism

- R: Randomisation
- T: Treatment
- C: ICE (treatment discontinuation)
- $\mathrm{Y_t}$: Outcome at time t
- U: Confounder (behavioural trait)
- $\mathrm{M_{y2}}$: A mechanism for assigning a value to Y.

A Blinded Randomised Experiment with ICEs

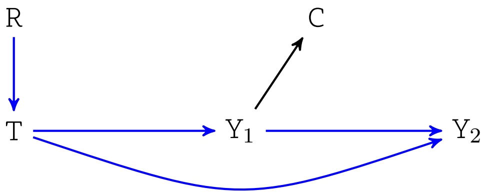


Figure 5. DAG of an ICE mechanism

- R: Randomisation
- T: Treatment
- C: ICE (treatment discontinuation)
- $\mathrm{Y_t}$: Outcome at time t
- U: Confounder (behavioural trait)
- $\mathrm{M_{y2}}$: A mechanism for assigning a value to Y.

A Blinded Randomised Experiment with ICEs

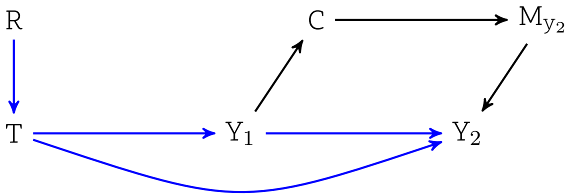


Figure 5. DAG of an ICE mechanism

- R: Randomisation
- T: Treatment
- C: ICE (treatment discontinuation)
- $\mathrm{Y_t}$: Outcome at time t
- U: Confounder (behavioural trait)
- $\mathrm{M_{y2}}$: A mechanism for assigning a value to Y.

A Blinded Randomised Experiment with ICEs

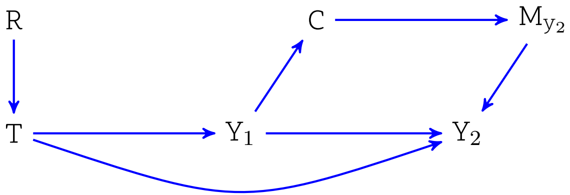


Figure 5. DAG of an ICE mechanism

- R: Randomisation
- T: Treatment
- C: ICE (treatment discontinuation)
- Y_t: Outcome at time t
- U: Confounder (behavioural trait)
- M_{y2}: A mechanism for assigning a value to Y.

A Blinded Randomised Experiment with ICEs

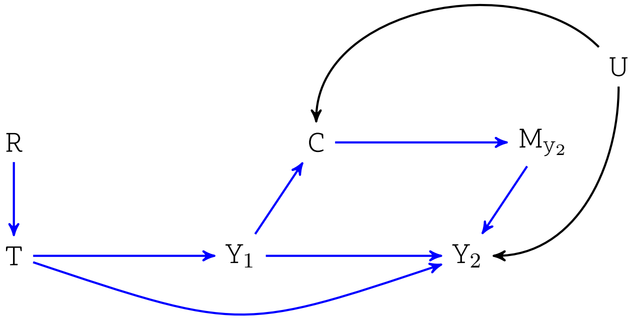


Figure 5. DAG of an ICE mechanism

- R: Randomisation
- T: Treatment
- C: ICE (treatment discontinuation)
- $\mathrm{Y_t}$: Outcome at time t
- U: Confounder (behavioural trait)
- $\mathrm{M_{y2}}$: A mechanism for assigning a value to Y.

A Blinded Randomised Experiment with ICEs

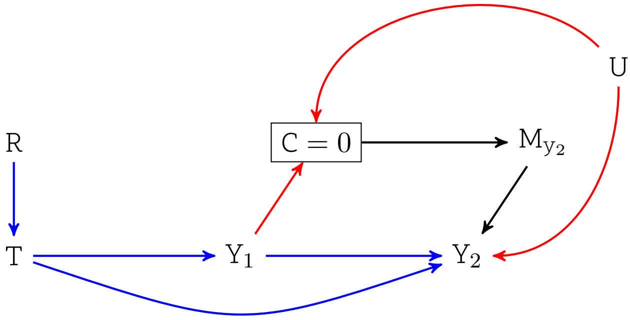


Figure 5. DAG of an ICE mechanism

- R: Randomisation
- T: Treatment
- C: ICE (treatment discontinuation)
- $\mathrm{Y_t}$: Outcome at time t
- U: Confounder (behavioural trait)
- $\mathrm{M_{y2}}$: A mechanism for assigning a value to Y.

A Blinded Randomised Experiment with Landmark Analysis

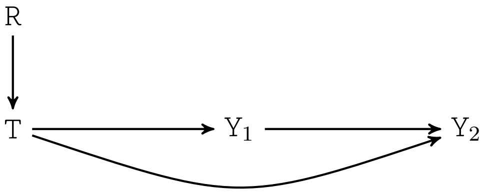


Figure 6. DAG of a landmark analysis in a RCT

- R: Randomisation
- T: Treatment
- $\mathrm{Y_t}$: Outcome at time t (survival)
- U: Confounder (patient global health)

A Blinded Randomised Experiment with Landmark Analysis

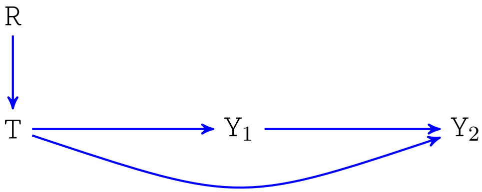


Figure 6. DAG of a landmark analysis in a RCT

- R: Randomisation
- T: Treatment
- Y_{t}: Outcome at time t (survival)
- U: Confounder (patient global health)

A Blinded Randomised Experiment with Landmark Analysis

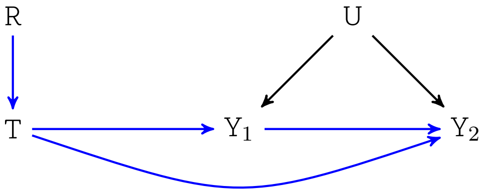


Figure 6. DAG of a landmark analysis in a RCT

- R: Randomisation
- T: Treatment
- Y_t: Outcome at time t (survival)
- U: Confounder (patient global health)

A Blinded Randomised Experiment with Landmark Analysis

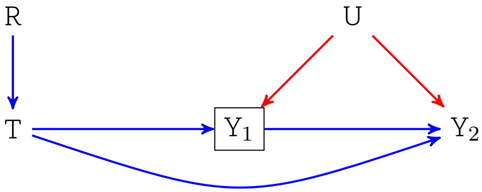


Figure 6. DAG of a landmark analysis in a RCT

- R: Randomisation
- T: Treatment
- Y_t: Outcome at time t (survival)
- U: Confounder (patient global health)

Contact Information

joshua.entrop@ki.se

joshua-entrop.com

Link to the slides and materials