Twin Networks

Alessio Zanga[1, *]

December 05, 2025

[1]Models and Algorithms for Data and Text Mining (MADLab), University of Milano-Bicocca, Milan, Italy

[*]Corresponding author: alessio.zanga@unimib.it

Attribution Analysis

Definition of Attribution

- Attribution quantifies **how much responsibility** a specific cause X carries for an observed outcome Y.
- It characterizes, in probabilistic terms, whether the outcome would still have happened without the cause (**necessity**), whether the cause would have produced the outcome on its own (**sufficiency**), or whether the cause was both required and adequate to bring about the outcome (**necessity & sufficiency**).

Attributional Queries (1/2)

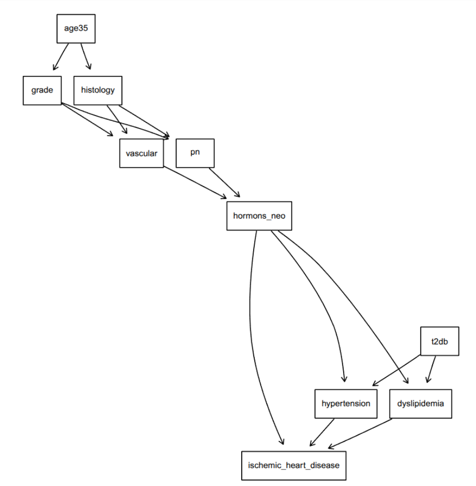
Probability of ...	**Definition**
Necessity (PN)	Probability that the outcome would not have occurred **without the treatment**, given that both the treatment and the outcome actually occurred.
Sufficiency (PS)	Probability that the outcome would have occurred **had the treatment been given**, given that neither the treatment nor the outcome actually occurred.
Necessity and Sufficiency (PNS)	Probability that the treatment is **both necessary and sufficient** to produce the outcome.

Attributional Queries (2/2)

Probability of ...	Formula
Necessity (PN)	$PN = P(Y_0 = 0 \mid X = 1, Y = 1)$
Sufficiency (PS)	$PS = P(Y_1 = 1 \mid X = 0, Y = 0)$
Necessity and Sufficiency (PNS)	$PNS = P(Y_1 = 1, Y_0 = 0)$

Counterfactual Explanations of Cardiovascular Risk (1/2)

- We investigate the role of the ovarian suppression therapy on ischemic heart disease and their potential precursors (i.e. hypertension and dyslipidemia).
- A causal network model that leverage on observational data on 1-year AYA BC survivors living the Lombardy region in Italy.



A. Bernasconi, et al. "On Counterfactual Explanations of Cardiovascular Risk [...] ." CEUR Workshop Proceedings. Vol. 3880. 2024.

Counterfactual Explanations of Cardiovascular Risk (2/2)

- $\mathcal{Q}$: Is ovarian suppression **a necessary and sufficient cause** for ischemic heart disease?

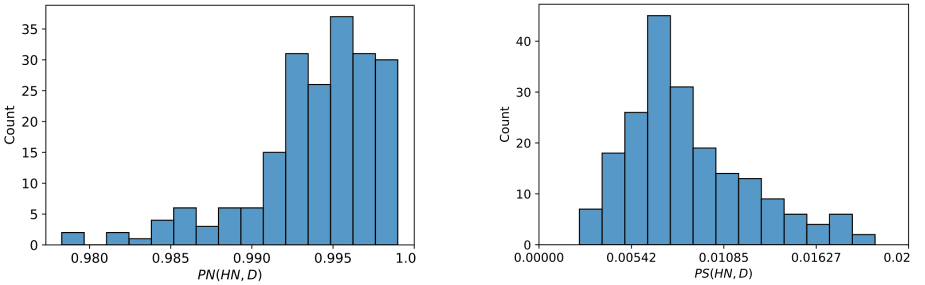


Fig. 1 Values of the probability of necessity (left) and probability of sufficiency (right) of variable `[hormons_neo]` for `[ischemic_heart_disease]`

Mediation Analysis

Definition of Mediation (1/2)

- Mediation analysis is the study of how a treatment X influences an outcome Y through one or more intermediate variables, called mediators M.
- **It decomposes the total effect** of X on Y into:
  - Direct effects - the effect that **occurs independently of the mediator**.
  - Indirect effects - the effect that **operates through the mediator**.
- Mediation analysis allows us to understand the **pathways by which a cause produces an effect**, quantify their contributions, and identify intervention targets.

Definition of Mediation (2/2)

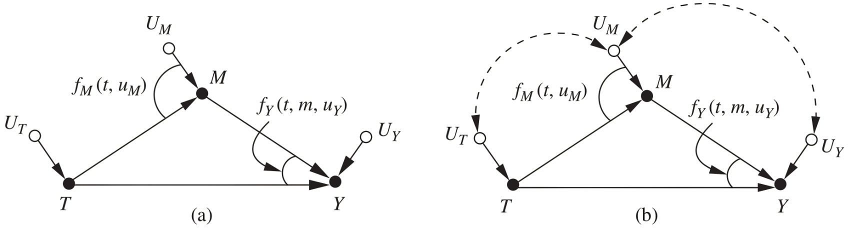


Figure 4.6 (a) The basic nonparametric mediation model, with no confounding. (b) A confounded mediation model in which dependence exists between U_M and (U_T, U_Y)

Mediation Queries (1/4)

- **Total Effect (TE)**
  - **Definition:** TE measures the expected increase in the outcome Y as the treatment changes from $X = 0$ to $X = 1$.
  - **Example:** If X indicates whether a patient receives a new drug (1) or a placebo (0), and Y is their recovery score, then the TE is the **expected improvement in recovery** caused by giving the drug instead of the placebo.
  - **Formula:**

$$\begin{aligned} \mathrm{TE}(X, Y) &= \mathbb{E}[Y_1 - Y_0] = \\ =&\mathbb{E}[Y_{\mathrm{do}(X=1)} - Y_{\mathrm{do}(X=0)}] = \\ =&\mathbb{E}[Y \mid \mathrm{do}(X = 1)] - \mathbb{E}[Y \mid \mathrm{do}(X = 0)] \end{aligned}$$

Mediation Queries (2/4)

- **Controlled Direct Effect (CDE)**
  - **Definition:** CDE measures the expected increase in the outcome Y as the treatment changes from $X = 0$ to $X = 1$, while the mediator is set to a specified level $M = m$ *over the entire population*.
  - **Example:** If X is a job-training program, M is the number of training hours, and Y is the earnings, the CDE at m is the **expected change in earnings** if everyone were to complete exactly m hours of training, compared to not doing training.
  - **Formula:**

$$
\begin{aligned}
\mathrm{CDE}(X, Y, M) &= \mathbb{E}[Y_{1,m} - Y_{0,m}] = \\
&= \mathbb{E}[Y_{\mathrm{do}(X=1,M=m)} - Y_{\mathrm{do}(X=0,M=m)}] = \\
&= \mathbb{E}[Y \mid \mathrm{do}(X = 1, M = m)] - \mathbb{E}[Y \mid \mathrm{do}(X = 0, M = m)]
\end{aligned}
$$

Mediation Queries (3/4)

- **Natural Direct Effect (NDE)**
  - **Definition:** NDE measures the expected increase in the outcome Y as the treatment changes from $X = 0$ to $X = 1$, while the mediator is set to whatever value it was prior to the change under $X = 0$.
  - **Example:** Let X be applying a new fertilizer (1) versus not applying it (0), M be the soil-microbe activity level the field would have had under no fertilizer, and Y be crop yield. The NDE captures **how much yield would rise from using the fertilizer** if the soil-microbe activity were held fixed at the level it would have naturally taken without the fertilizer.
  - **Formula:**

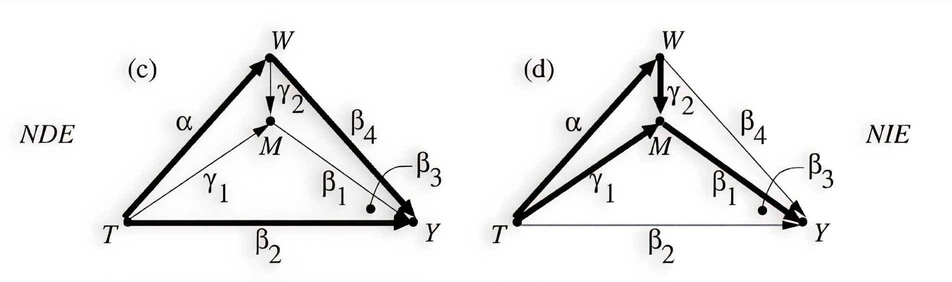
$$\mathrm{NDE}(X, Y, M) = \mathbb{E}[Y_{1,M_0} - Y_{0,M_0}]$$

Mediation Queries (4/4)

- **Natural Indirect Effect (NIE)**
  - **Definition:** NIE measures the expected increase in the outcome Y when the treatment is held constant, at $X = 0$, and M changes to whatever value it was under $X = 1$.
  - **Example:** If X is applying an irrigation system, M is the resulting soil moisture level, and Y is crop yield, the NIE captures **how much yield would rise purely because soil moisture** is moved to the level it would have had with irrigation - even though the irrigation itself is not applied.
  - **Formula:**

$$\mathrm{NIE}(X, Y, M) = \mathbb{E}[Y_{0,M_1} - Y_{0,M_0}]$$

Visualizing the Natural Effects



Social Network Properties and Metabolic Syndrome (1/4)

- Metabolic syndrome is a **constellation of conditions** that occur together and increase the risk of cardiovascular diseases.
- Includes obesity, low high-density lipoprotein, high triglyceride, hypertension and hyperglycemia.

Significance of this study

What is already known about this subject?

- Research has shown that social network properties are associated with metabolic syndrome, which is a major risk factor for cardiovascular diseases.

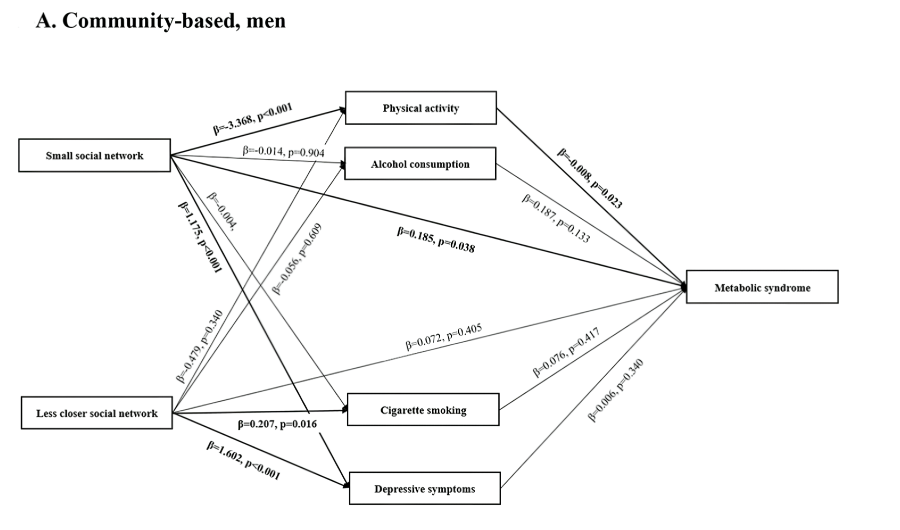
What are the new findings?

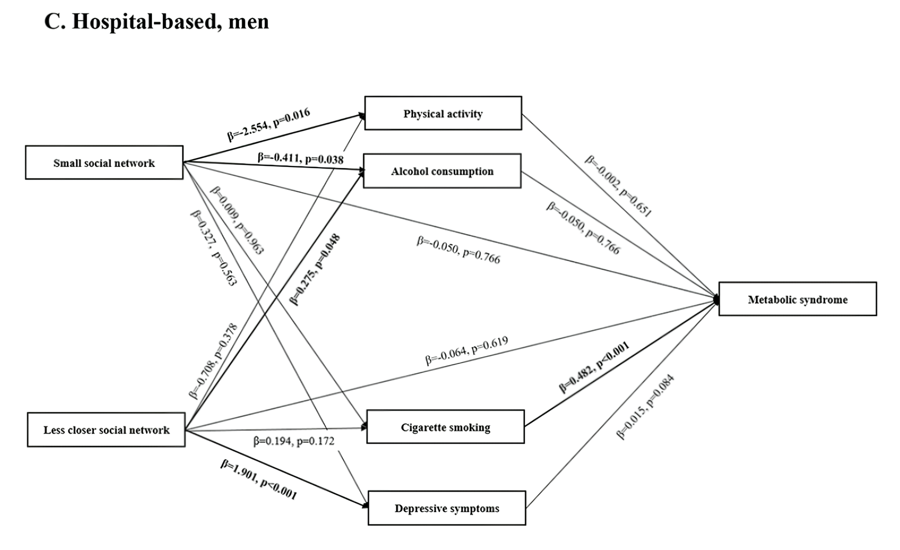
- We found that a small social network is positively associated with metabolic syndrome in Koreans.
- The association between social networks and metabolic syndrome is partially mediated by decreased physical activity, especially in older individuals.

How might these results change the focus of research or clinical practice?

- Our study suggests that developing strategies to encourage physical activity in socially isolated individuals could help prevent metabolic syndrome.

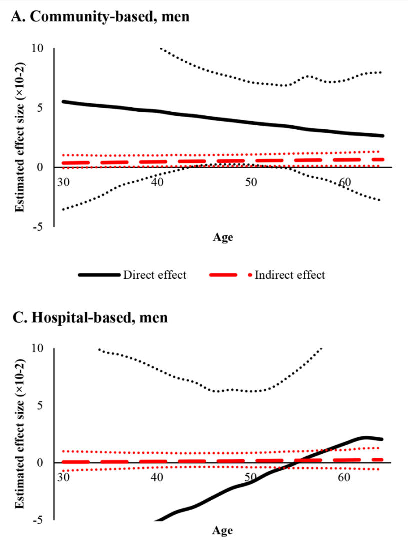
K., Kwanghyun, et al. "Associations between social network properties and metabolic syndrome [...]." BMJ Open Diabetes Research & Care (2020).


A. Community-based, men
Physical activity
Alcohol consumption
Small social network
Metabolic syndrome
Less closer social network
Cigarette smoking
Depressive symptoms
β=-3.368, p<0.001
β=-0.014, p=0.904
β=-0.004,
β=1.175, p<0.001
β=0.185, p=0.038
β=-0.008, p=0.023
β=0.187, p=0.133
β=-0.056, p=0.609
β=-0.479, p=0.340
β=0.072, p=0.405
β=0.076, p=0.417
β=0.006, p=0.340
β=0.207, p=0.016
β=1.602, p<0.001


C. Hospital-based, men
Physical activity
Alcohol consumption
Small social network
Metabolic syndrome
Cigarette smoking
Less closer social network
Depressive symptoms
β=-2.554, p=0.016
β=-0.411, p=0.038
β=0.009, p=0.963
β=0.327, p=0.563
β=-0.050, p=0.766
β=0.275, p=0.048
β=-0.002, p=0.651
β=-0.050, p=0.766
β=-0.708, p=0.378
β=-0.064, p=0.619
β=0.482, p<0.001
β=0.015, p=0.084
β=0.194, p=0.172
β=1.901, p<0.001

Social Network Properties and Metabolic Syndrome (4/4)

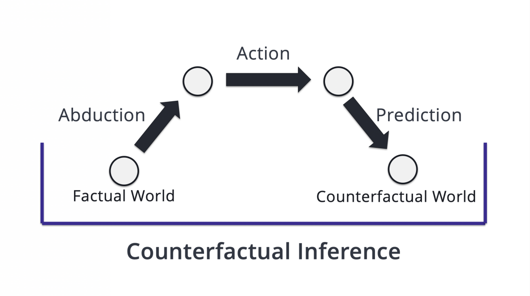
- Age-specific direct effect of size of social network and indirect effect through physical activity.
- In community-based population, age-specific indirect effect of network size through physical activity was larger in older age.
- In contrast, **no significant trends in indirect effects were found in hospital-based population**.



Twin Networks

Counterfactual Estimation

- Causal model M with noise $P(U)$, the counterfactual $P(Y_x \mid e)$ is given by:
  1. **Abduction** - Update the distribution with evidence e to obtain $P(U \mid e)$
  2. **Action** - Apply the action $do(x)$ to obtain M_x
  3. **Prediction** - Use the intervened model M_x with updated noise distribution $P(U \mid e)$ to compute the probability of Y, the consequence of the counterfactual.

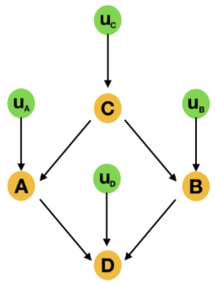


The Problem

- The first step, **requires a large amount of computational resources**.
- Moreover, as **this step has to be repeated** for every new counterfactual query, such computational resources are continually required.
- In the case of k binary variables and exact inference, $P(U \mid e)$ grows with 2^{k-1}, which **for $\mathbf{k = 30}$ is $\sim \mathbf{500}$ million**.

Definition of Twin Network

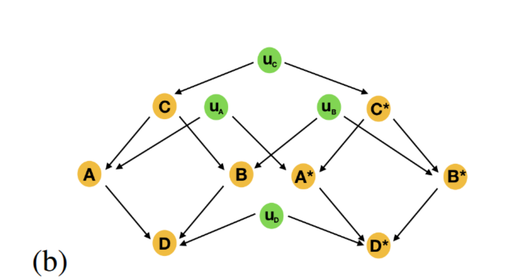
- $\mathcal{Q}$: Counterfactual statement to be computed $P(D_{A=0} = 0 \mid D = 1)$.
- The probability that D would have been 0 had A been forced to be 0, given that $D = 1$ was observed.
- A **twin network** employs two inter-linked networks, one representing the real factual world and the other the counterfactual world being queried.



(a)

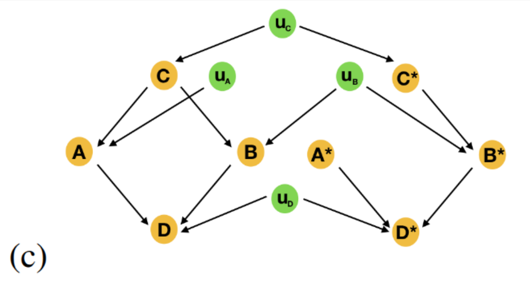
Querying of Twin Network (1/4)

- The twin network approach to this problem first **constructs the linked factual and counterfactual networks** depicted in (b), with counterfactual nodes represented with a superscript $*$.



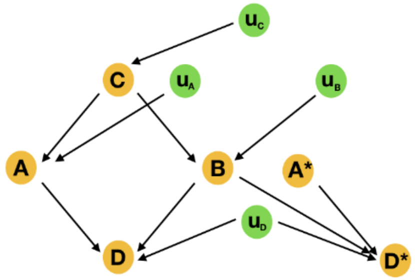
Querying of Twin Network (2/4)

- The intervention $do(A^* = 0)$ is performed in the counterfactual network.
- **All ingoing arrows** from the parents of A^* to A^* itself **are removed** and A^* is **set to 0**. This is depicted in (c).



Querying of Twin Network (3/4)

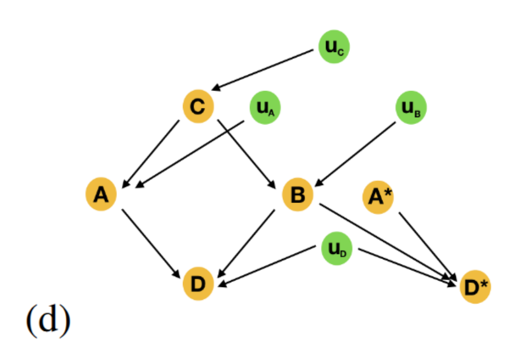
- The counterfactual query is reduced to the **conditional probability in the twin network**: $P(D^* = 0 | D = 1)$, which can be computed using standard belief propagation techniques.



(d)

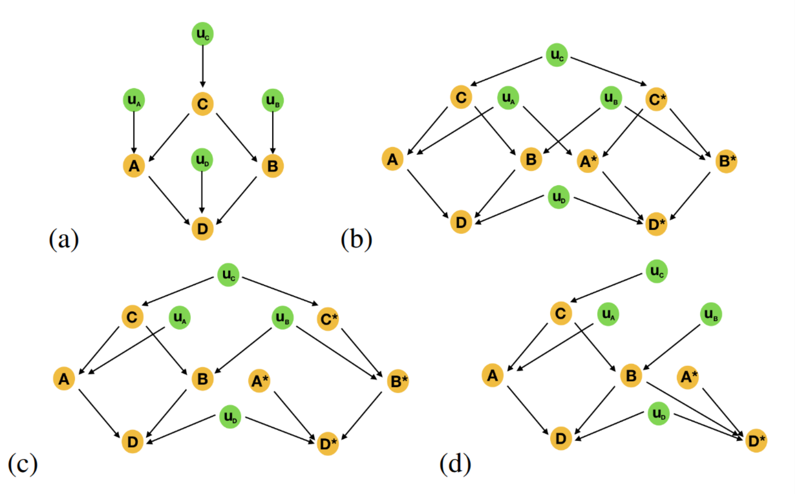
Querying of Twin Network (4/4)

- This allows the counterfactual query to be computed **without the need to store the joint distribution** over U_C and U_A (conditioning on $D = 1$ induces correlations between them).



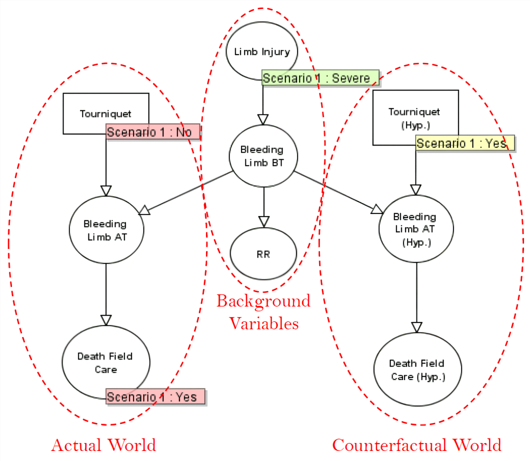
The Big Picture

- (a) Causal Network, (b) Twin Network, (c) Intervention, (d) Node-merging.



Counterfactual Reasoning [...] as a Healthcare Governance Tool (1/3)

- **Defence Medical Services (DMS)** mortality and morbidity (M&M) review meeting, measures the performance of the UK deployed clinical services and to provide beneficial feedback.
- Predict **the risk of an injured soldier dying during field care**.

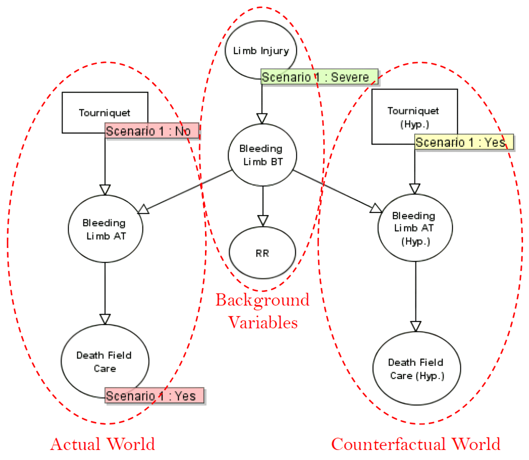


Kyrimi, E., et al. "Counterfactual reasoning using [...] as a healthcare governance tool." International journal of medical informatics 193 (2025).

Counterfactual Reasoning [...] as a Healthcare Governance Tool (2/3)

- In the actual world, we observe that although the soldier **had a severe limb injury**, a tourniquet[a] was not applied, and the soldier died.
- Imagining a counterfactual world, we would like to know whether the soldier would have survived if we had applied a tourniquet.

[a] *ITA: Laccio emostatico.*



Counterfactual Reasoning [...] as a Healthcare Governance Tool (3/3)