MATH464/564: Practical 2a: Directed acyclic graphs and confounding

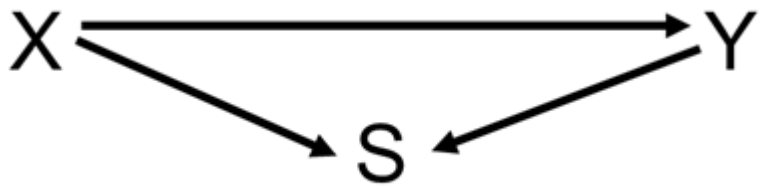
Tom Palmer

19/20

Question 1

In the following DAG:

- the outcome Y is an indicator of whether the baby had neural tube defects
- the exposure E indicates whether the mother took folic acid supplementation in early pregnancy
- the third variable S indicates whether the pregnancy ends in a still birth or therapeutic abortion



It was noticed that controlling for S in the model leads to a substantial change in the estimated association between X and Y. Should we condition on S? Give your reasons.

Question 2

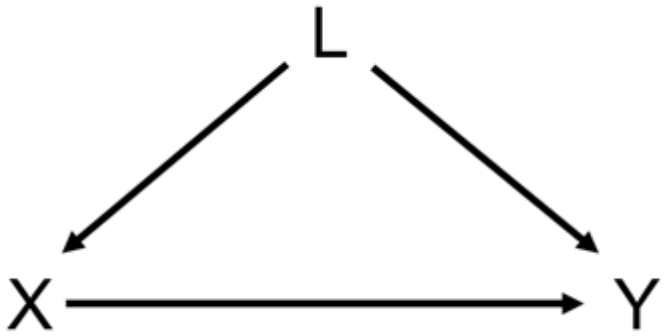
In this question we are attempting to estimate the effect of X on Y.

a) Write down the (old) definition of a confounder

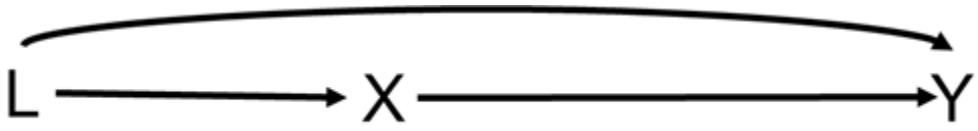
b) Write down Pearl's definition of a confounder

c) For each of the following DAGs assess whether L is confounder using both the old and new definitions.

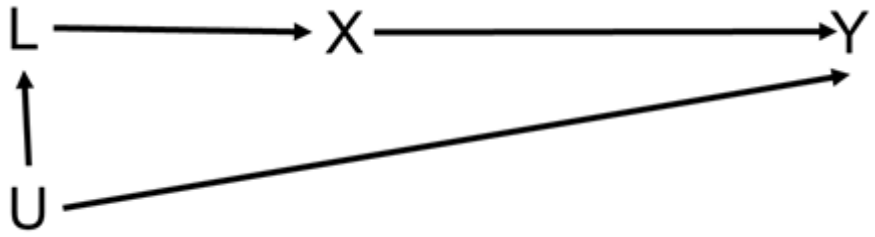
DAG A



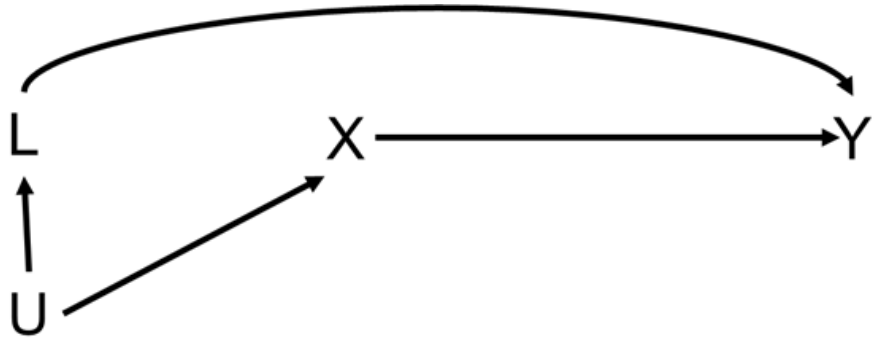
DAG B



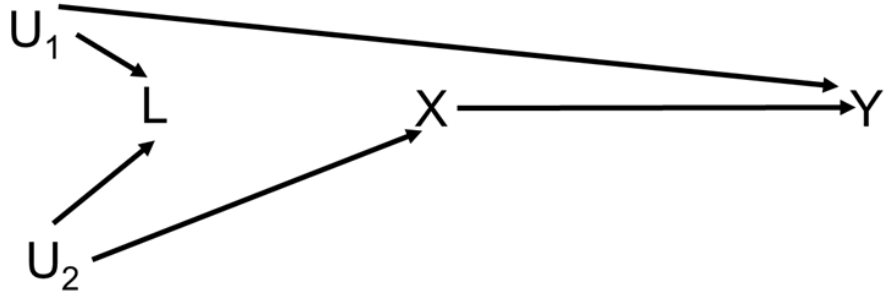
DAG C



DAG D



DAG E



Question 3

The fetal origins hypothesis suggests that there are associations between prenatal growth and adult heart disease.

Huxley et al., Unravelling the fetal origins hypothesis: is there really an inverse association between birthweight and subsequent blood pressure? Lancet, 2002, 360, 9334, 659–665, https://doi.org/10.1016/S0140-6736(02)09834-3, reviewed 55 studies that reported associations between birthweight, the exposure, and later life systolic blood pressure, the outcome. The table below shows what potential confounding factors the authors adjusted for.

Potential confounding factor	Number of studies adjusting for factor
Current weight	49
Sex	48
Height	13
Parental socioeconomic status	7
Current socioeconomic status	2
Parental blood pressure	9
Alcohol consumption	3
Race	6
Gestational age	8

Other factors adjusted for (and number of studies that adjusted for each factor): ambient temperature or exercise (4 studies); sphygmomanometer cuff size (3 studies); amount of television watched, anticipated venepuncture, heart rate, maternal body-mass index, parity, person who measured blood pressure, Tanner's stage of puberty, or town (2 studies); and Apgar score, birth rank, calcium in pregnancy, father's height, maternal age, maternal haemoglobin, maternal oedema, cigarette tar dose, or time of day (1 study).

Table 2: **Adjustment for potential confounding factors in the 55 studies that reported regression coefficients for the association between birthweight and subsequent blood pressure**

As you can see the vast majority of these associations were adjusted for current weight. In the Results section the authors commented as follows.

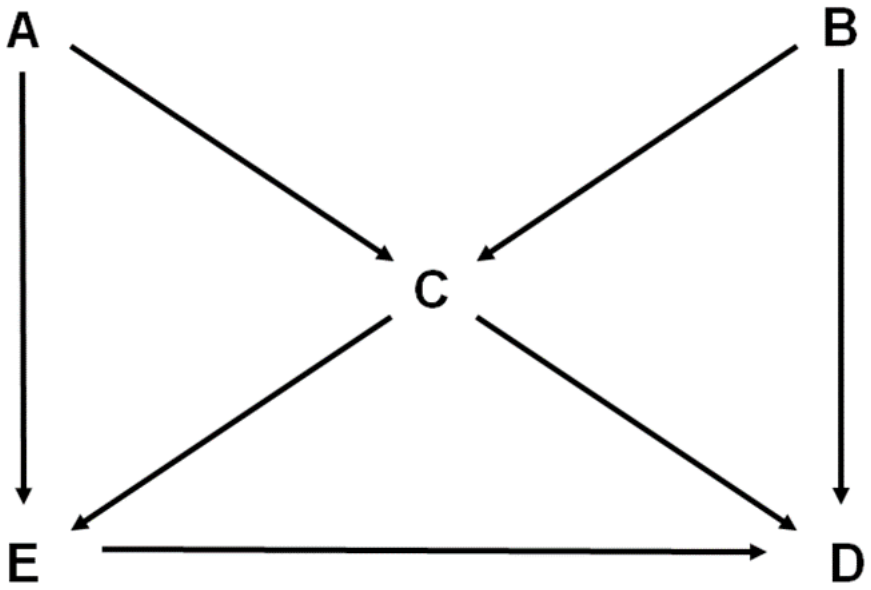
> birthweight. Regression coefficients without adjustment for current size have now been obtained for 12 of the 17 studies that included more than 1000 individuals (see Acknowledgments), representing nearly 350 000 of the 380 000 individuals contributing to the previous quantitative estimates. In those studies, removal of the adjustment for current weight produced about a halving in the inverse-variance-weighted estimate for the association between birthweight[22] and subsequent systolic blood pressure: –0·6 mm Hg/kg with adjustment for current weight reduced to –0·4 mm Hg/kg without adjustment. As

So, removing the adjustment for current weight, in 12 studies, halved the magnitude of the pooled association. The pooled association with adjustment was -0.6 mmHg/kg; and without adjustment was -0.4 mmHg/kg.

Draw a possible DAG for the relationship between birthweight, current weight, and current SBP.

Question 4

The following DAG describes a model between an exposure (E), a disease outcome (D), and three other variables (A, B, and C).



a) List all the backdoor paths between E and D.

b) With respect to A and B how would you describe C?

c) If we fitted a regression model using B as the outcome regressed on covariates A and C, what could be the problem?

d) Suppose we want to estimate the effect of E on D, is there one best model to fit to estimate this association? Write down this best model/s.

e) Verify your answers using the dagitty package in R.