Solutions 5: Confounding and causality

Question 1: Diabetes and CHD

Question 1a:

Diabetes was associated with a significantly increased rate of CHD death during 20 years of follow-up in both men and women. Women with diabetes had a 3.8-fold increased rate of CHD death and men with diabetes had a 2.1-fold increased rate of CHD death, compared to their counterparts without diabetes.

Question 1b:

All of these factors are associated with diabetes. In order to be considered as potential confounders we need to know if they were also independently associated with CHD mortality. It is likely that factors such as smoking and age are risk factors for CHD mortality and not on the causal pathway between diabetes and CHD mortality, and therefore we would consider them as potential confounders. The other factors are also likely to be associated with CHD mortality, but we would want more subject specific knowledge to have a better idea as to whether they are independent risk factors or on the causal pathway between diabetes and CHD mortality. (In fact, the authors of this research consider them as potential confounders.)

Question 1c:

The researchers can stratify their analysis by individual confounders, and then summarise the stratum-specific measure of effect using techniques such as Mantel Haenszel or Standardisation to adjust for these confounders. If they want to control for multiple confounders simultaneously (which is usually the case), they can use multivariate modelling techniques to do this.

Question 1d:

If no information is collected on potential confounders, it is not possible to control for them in subsequent analyses. Information MUST be collected on potential confounders in order to assess their roles as alternative explanations for observed exposure/disease relationships.

Question 1e:

(i) There is evidence of partial confounding by age. The age-adjusted RR (5.2) is weaker than the crude RR (8.6), indicating that once age differences between diabetics and non-diabetics are taken into account, the association between diabetes and CHD mortality is not as strong as it previously appeared. However, the association was not completely confounded/explained by age. The age-adjusted RR is still far from 1 (with 95% confidence that the true population parameter is between 2.2 and 6.6), meaning that even once differences in age are taken into account, there is strong evidence of an association between diabetes and CHD mortality.

[For those interested in CI: note also that the CI for the adjusted RR is narrower. CI indicates the how precisely the sample estimate corresponds to the true effect

(RR in this case) in the population. The width of the CI is affected by 1. sample size, and 2. variability in the population. Sample size has not become smaller in this case. If adjusting leads to more precise estimate in each stratum and in turn in the adjusted RR (because the extent to which data in each stratum varies is less than the overall variability), the standard error would be smaller and therefore CI narrower.]

(ii) Older people are at higher risk for CHD mortality. Age is a risk factor for diabetes. And age is not on the causal pathway between diabetes and CHD mortality. Thus, age fulfils the criteria for being a confounder.

After adjustment for age, we can conclude that the observed RR of 5.2 is a better estimate of the true effect of diabetes on CHD mortality, whereas the unadjusted RR of 8.6 was a distorted estimate.

(iii) After controlling for these additional variables, the association between diabetes and CHD mortality is weaker than the crude RR and the RR which was only adjusted for age. The change in RR is consistent with the inclusion of confounding variables in the regression model. Before making this conclusion, we must make sure that these 'third variables' fulfil the three criteria for being a potential confounder (as opposed to being a mediating factor). And then we can state that the multivariable-adjusted RR of 3.8 is an even better estimate of the true effect of diabetes on CHD mortality than the age-adjusted estimate of 5.2.

(iv) Diabetes still appears to be a risk factor for CHD mortality. After adjustment for age, smoking, BMI, etc, women who had diabetes are 3.8 times more likely to die from CHD over 20 years, than women did not have diabetes. It is possible that there is confounding by other variables which were not measured/taken into account in the analysis. Examples of such variables we might want to consider are socioeconomic status, diet, etc.

Question 2: Cholera and water

1.

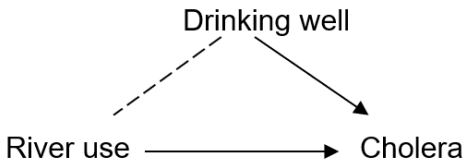
- Comparison groups: people who use river water for washing verses people who do not use the river for washing.
- Outcome: Incidence rate of cholera.
- H_0: There is no associate between using the river for washing and cholera. (In other words: the incidence rate of cholera among river users = incidence rate of cholera among nonusers; RR=1.0).

2. $Incidence\ rate\ ratio = \frac{210/30000}{70/15000} = 1.50$

Healthy adult residents who use river water for washing had 1.50 times (one and a half times) the rate of cholera when compared to healthy adult residents who did not use river water for washing.

3. The epidemiologists already knew the drinking well was contaminated, so it was an established risk factor for cholera. It was possible that there was some association between well use and river use, perhaps relating to the socioeconomic status of participants and/or their place of residence. And it's not plausible that well usage was on

a causal pathway between river use and cholera. So, there's an arguable case that 'drinking well' is a confounder.



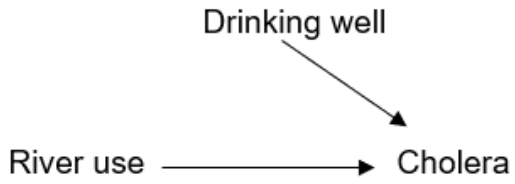
4.

- Rate ratio for the river water-cholera association among well users = 1.50
- Among people who drank water from the contaminated well, the rate of cholera for river water users was 1.50 times the rate of cholera for river non users.
- Rate ratio for the river water-cholera association among non-users = 1.50
- Among people who did not drink water from the contaminated well, the rate of cholera for river water users was 1.50 times the rate of cholera for river non users.
- There is a complicated formula for calculating the weighted average of stratum-specific RRs. Thankfully in this case the average of 1.50 and 1.50 has to be 1.50! The well-adjusted RR is 1.50.

After adjusting for drinking well use, the rate of cholera among river water users was 1.50 times (50% higher) the rate of cholera among non-users of river water.

5. When the crude RR and the adjusted RR are essentially the same, it suggests that that the variable used for stratification (in this case drinking well) was not a confounder. To establish whether the river is a risk factor for cholera, the epidemiologists still need to consider whether there are other confounding factors, whether there is bias in the selection of participants, whether there is bias in the measures, and whether there was adequate statistical power to detect an association.

   The new diagram looks like this:



   Which means that 'drinking well' is probably an independent risk factor for cholera, not a confounder.

Question 3: Causation

a) Vitamin A supplementation and maternal mortality. Even though there was strong evidence of a reduction in the risk of maternal mortality in the supplemented group, and a well conducted RCT design, this is likely to be a chance finding, especially since the risk reduction was mainly seen in the injuries and miscellaneous group and there is no biologically plausible reason for arguing that supplementation could reduce these deaths. Subsequent trials from Ghana and Bangladesh showed no association between supplementation with vitamin A and maternal mortality.

b) Underweight and mortality. The increased risk of mortality in the underweight group is likely to be due to confounding by underlying health conditions and smoking. Studies have shown that after adjusting for smoking and underlying health conditions, the risk of mortality in the underweight group is lower than in the normal weight group.

c) Amyl nitrite and Kaposi's sarcoma. The association between Amyl nitrite and Karposi's sarcoma was confounded by HIV. Amyl nitrite is a recreational inhalant which at that time was commonly used by men who were having sex with other men. Unprotected sexual contact is a transmission risk for the Human Immunodeficiency Virus (which was yet to be discovered, and is now known to be the causative agent for AIDS).

d) MMR and autism. The age at which MMR vaccination is given is the same age at which behavioural problems become apparent, which would make the two factors appear to be associated. The sample size is very small and there is no appropriate control group who were unexposed. This is the study that triggered the mass panic in the UK about the link between autism and MMR vaccination. Further issues, including the fact the study had not been approved by the local ethics committee, meant that the paper was subsequently retracted by the editors of the journal, twelve years after publication.