

Confounding by Population Stratification A Guided Reading in Genetic Epidemiology

Answer Key

Question 1: Using only the information in Table 2 of the Knowler paper, what might you conclude about the GM marker and risk of diabetes?

YOUR ANSWER:

Using either the prevalence ratio of 0.27 (95% CI: 0.18-0.40) or the odds ratio of 0.21 (95% CI: 0.13-0.32) the Gm marker is inversely associated with type II diabetes in this population. This could be erroneously interpreted to indicate that the absence of the Gm marker is a risk factor for diabetes.

Question 2: Using only the information in Table 4 of the Knowler paper, quantify the association between the Gm marker and type II diabetes. What might you conclude from this result about the GM marker and risk of type II diabetes?

YOUR ANSWER:

Using either the prevalence ratio of 0.98 (95% CI: 0.66-1.46) or the odds ratio of 0.95 (95% CI: 0.33-2.78) there is no evidence for association between the Gm marker and risk of type II diabetes. Presence or absence of the Gm marker does not appear to be related to increased risk of diabetes in full-heritage Pima-Papago Indians 35 years of age or older.

Question 3: What is confounding?

YOUR ANSWER:

Confounding is a distortion of the level of association between an exposure and disease. Confounding results from the effects of a third variable (the confounder), which is associated with both the exposure and the disease.

Question 4:

Looking at Figure 3 of the Knowler article, what does the left hand panel indicate about the association between level of American Indian heritage and outcome (type II diabetes)? What does the right hand panel indicate about the association between level of American g heritage and exposure (Gm haplotype)? Are the conditions for confounding by heritage met in this situation? Why or why not?

YOUR ANSWER:

The left hand panel shows that level of Indian heritage is associated with outcome. The prevalence of type II diabetes increases with increasing level of Indian heritage.

The right hand panel shows that level of Indian heritage is associated with disease. The prevalence of the Gm Marker decreases with increasing level of Indian heritage.

Indian heritage fits the classic definition of a counfounder. Indian heritage likely confounds the association between Gm Marker and diabetes.

Question 5: Using this information, what is the age-adjusted prevalence ratio for individuals of full American Indian heritage? 50% American Indian heritage? Full Caucasian heritage? What would you expect to be the approximate prevalence ratio if you adjusted for both age and American Indian heritage (you do not need to do actual calculations, just give an approximate answer)?

YOUR ANSWER:

Full Pima-Papago: Age-adjusted prevalence ratio= 0.91
50% Pima-Papago: Age-adjusted prevalence ratio= 0.98
Full Cacuasian: Age-adjusted prevalence ratio= 0.89

Adjusting for both age and degree of American Indian heritage would result in a prevalence ratio that was a weighted average of the three strata specific prevalence ratios. Therefore, I would expect the age and heritage adjusted prevalence ratio to be ~0.95. We have not done a significance test, but this prevalence ratio is close to one, indicating that after adjustment for age and admixture, there is not a strong association between the Gm marker and type II diabetes.

Question 6: Overall, how do you interpret the findings of this study? What are the implications of the findings (if any) for genetic association studies? What are the implications of the findings (if any) for public health?

YOUR ANSWER:

This study demonstrates that genetic admixture (a form of population stratification) can confound the results of a genetic association study. When genetic admixture is controlled for, there is not an association between the Gm marker and type II diabetes.

This study shows the importance of considering population stratification in genetic association studies. Failure to properly account for population stratification could result in spurious associations.

Research in public health should continue to collect information on family history, including ethnicity. This information can be used to insure that genetic studies properly control for potential population stratification.