Comments on Biostatistics 208 Lab #4 1/30/20

Confounding

Question 1. What is the interpretation of the unadjusted and adjusted exponentiated regression coefficients for BMI, also the `lincom` and nlcom results?

As discussed last week, when the outcome is natural log-transformed, exp(βk) estimates the relative change in the mean of the untransformed outcome for each k-unit increase in the predictor, and 100(exp(βk) – 1) estimates the corresponding percent increase. With the `eform("exp(beta)")` option in regress, the coefficient estimates are labeled in the Stata output as `exp(beta)`.

Thus in the unadjusted analysis, the estimate in the regress output shows that each kg/m2 increase in BMI is associated with a 1.002-fold increase in creatinine (95% CI 1.0007-1.0036, $p = 0.004$), after rounding the results a bit. (See output in log file provided on course web site.)

After adjustment, a 1 kg/m2 increase only predicts only 1.0015-fold higher creatinine (95% CI 1.0001-1.0030, $p = 0.037$). (See output in log.)

Mediation

It may be helpful to use a DAG like the one included on the last page in thinking about mediation in this context.

Question 2. Is BMI an independent predictor of TG levels? Or in other words, is there evidence that exposure affects the proposed mediator?

The `nlcom` result following the fit of model 1 (below) show that each 5 kg/m2 increase in BMI is independently associated with an almost 7.0% increase in TG levels ($p < 0.0005$). Only a moderate effect, but highly statistically significant. So there is evidence for the first link the proposed mediating pathway.

Question 3. Are TG levels an independent predictor of creatinine levels? What is the interpretation of the `lincom` and `nlcom` results?

From model 3 (see log file), there is little question that TG is independently associated with creatinine ($p < 0.0005$) after adjustment for BMI as well as the covariates. The effect is again not large: the lincom result shows that a 25% increase in TG is associated with a 1.011-fold increase in creatinine; the `nlcom` result translates this into a 1.1% increase in creatinine. So there is evidence for the second link in the proposed mediating pathway.

Question 4. Is there convincing evidence for an indirect pathway from BMI to TG to creatinine? Interpret the indirect effect estimate. Does this pathway make biological sense?

We've just seen that the adjusted associations reflecting the first (BMI→TG) and second (TG→creatinine) links in the proposed indirect pathway are both highly statistically significant. The result using the command

```
dis 100*(exp(5*link1*link2)-1)
```

following model 3 shows that the indirect effect via TG of a 5 kg/m2 increase in BMI is only to increase

creatinine levels by about 0.34%.

A more difficult question is whether this proposed causal pathway makes biological sense. The adverse effect of obesity on lipids is reasonable, but a nephrology fellow working with Eric Vittinghoff reports that the second link is controversial. The analysis would need to control for confounding a lot more persuasively than these models do for the pathway to be taken seriously.

Question 5. What is the interpretation of the BMI effect after adjustment for TG?

This estimates the direct effect of BMI on creatinine via other pathways. After adjustment for log-TG, the association of BMI with creatinine is weaker and no longer statistically significant, with lower confidence bound less than zero, so evidence for a direct effect via other pathways is weak.

Question 6. What percentage of the BMI effect on creatinine is explained by TG levels? What do you make of the upper bound of the confidence interval?

The estimated percentage of the BMI effect explained by TG is 44.3% (bias-corrected bootstrap percentile confidence interval 15.3-287%). Recall that the overall effect is the sum of the direct and indirect effects. We have fairly strong evidence that the indirect effect on BMI via TG is to increase creatinine levels. Thus the upper confidence bound is consistent with an inverse direct effect of BMI on creatinine levels, as reflected in the lower confidence bound for BMI in model 3. To put it another way, if PE>100%, then TG explains all of the overall effect of BMI on creatinine and then some.

Question7. How do the `medeff` results compare to the earlier ones?

The `medeff` estimates for the indirect (i.e., ACME; 0.00068), direct (0.00084), and total effects (0.00152), as well as PE (44.4%, after rescaling) agree closely with the first result for the indirect effect (0.00068), just below Model 3, the direct effect estimate shown in the Model 3 output (0.00085), the total effect estimate shown in the Model 2 output (0.00154), and the PE estimate we calculated (44.3%); moreover, the confidence intervals for the direct effects are consistent in spanning zero. As noted in lecture, the agreement is not perfect because `medeff` uses a simulation approach, but the differences are trivial. The `medeff` confidence interval for PE is somewhat narrower (20.1-246) than the bootstrap result (15.3-287) but they are qualitatively consistent, especially in having an upper bound far above 100%. Estimates of this ratio statistic are known to be imprecise.