Biostat/Epi 536 2024
HW 6

The Studies of Left Ventricular Dysfunction (SOLVD) trials were two large randomized placebo-controlled trials of the efficacy of enalapril, an angiotensin converting enzyme (ACE) inhibitor, for the treatment and prevention of congestive heart failure to improve survival in patients with a weak left ventricular ejection fraction [1]. The original analyses were based on a time-to-event outcome; however, for the purposes of this homework, we consider a simulated version of the data that was based on this study but uses a binary version of the death outcome (I_death, 1=yes, 0 =no) that ignores the differential length of follow-up. Covariates in this dataset include the treatment arm indicator (I_ace, 1=ACE Inhibitor, 0=Placebo control) and the baseline characteristics: sex (I_sex), age in years (age), left ejection fraction (ef), and an indicator of whether they had diabetes (I_dm). Please use the `sovld_simulated.csv` data set to answer the following set of questions.

1. Use logistic regression to assess the unadjusted effect of the ACE inhibitor in this trial on survival (i.e. on the death outcome). Provide the odds ratio (OR) and confidence interval.
2. Suppose in planning the study, the investigators hypothesized that age, sex, ejection fraction and diabetes were all potential precision variables. What would need to be true about the association with outcome and exposure for these variables to be precision variables? Does the data seem to support the idea that each of these variables might be useful as precision variables? State your reasoning.
3. Fit the model in 1, now adding in age, sex, ejection fraction and diabetes. Did the association between treatment and outcome appear to get stronger, weaker or stay the same compared to the unadjusted model in #1?
4. Fit a model to evaluate whether ejection fraction (as a continuous variable) modifies the treatment effect on survival. What do you conclude from this model and summarize your conclusion in a sentence suitable for a scientific publication.
5. Now suppose the investigators wanted to evaluate whether having low ejection fraction (indicated by ef<20, yes/no) modifies the treatment's effect on death. Fit a model to evaluate this question and state your conclusion, including your rationale.
6. Using the model in #5, provide the treatment odds ratio and 95% confidence interval for those with and without low ejection fraction (assuming constant values for any other variables that are in your model for this comparison).
7. Suppose the effect of the baseline ejection fraction on the death outcome is now of interest. Can you use the coefficients in the model you fit in question 5 to evaluate whether treatment modifies the effect of low ejection fraction on death? If yes, discuss what is your conclusion.
8. (optional) Investigate whether the modification of the treatment effect on the ejection fraction appears to be non-linear.

REFERENCE

[1] SOLVD Investigators. Studies of left ventricular dysfunction (SOLVD)—rationale, design and methods: two trials that evaluate the effect of enalapril in patients with reduced ejection fraction. *The American Journal of Cardiology*. 1990 Aug 1;66(3):315-22.