**BIOST 2049**

**Spring 2023**

**Research Article Dissection 3A – Splines**

**Due 2/22/23 by 11:59 via the Assignment Link posted to Canvas**

Please read the article “Total and Percent Free Prostate-Specific Antigen Levels among U.S. Men, 2001-2002” by Saraiya et al. found on Canvas and answer the questions below.

This is an individual assignment. You can discuss the article with your classmates, but the answers you submit should be your own.

Please be brief but write in complete sentences with appropriate interpretations (a yes or no answer will not get full credit). A few sentences is all that is necessary to answer each question.

1. What was the purpose of this study?

Detect the effect of thresholding measures of PSA on prostate cancer screening and determine whether or not to increase distribution of a more novel test. Both based on demographic characteristics.

1. Briefly describe the methods used in this paper. Notice how they mixed results with the methods? Results should always be separate from the statistical methods. For this question, only focus on the methods (last two paragraphs of this section).

They fit two linear regression models. One tested the relationship between age and total PSA. The other tested the relationship between age and percent free PSA. Both models controlled for race and assessed interactions between race and age. They log transformed the PSA response variables and used an F-statistic to test significance.

1. Explain the rationale for using splines in this analysis. How were the splines applied? Do you think it was appropriate to use splines in this analysis?

Splines derive non-linear features from the data distribution of age to place in their linear models. They used restricted cubic spline functions. According to their graphs, it does appear that it was appropriate to use splines, since the Percent Free PSA by Age curve appears nonlinear, and since the age\*race by Total PSA curve appears nonlinear for certain races. However, they could have tested other models that are less complex too for the sake of comparison.

1. Explain the rationale for using an interaction between age and race/ethnicity. Was the interaction justified?

The curves fit between age and PSA measures appear to differ by race/ethnicity

1. Give your interpretation of the patterns in Figures 1 and 2. Do these graphs support the main conclusion?

Yes, measures of PSA appear to differ by age and this relationship appears to differ by race.