**Exam 2 – Fall 2022**

**Problem #1:**

Smokers were randomized, in an 8-week trial, to either smoking the same brand (control group; arm #5) or switching to e-cigarettes (arm #6); File: E-CIG-2-E-22. Outcome variables are final NNAL and TNE measures (NNAL\_vt4 and TNE\_vt4 of visit 4) but they are not the targets of the problem. In this problem, we do not study the outcomes; we focus on the baseline measures (NNAL\_vt0 and TNE\_vt0). Define:

X1 = log (baseline measure NNAL) and

X2 = log (baseline measure TNE)

The aim here is to is to see if the two arms 5 and 6 have similar values of X1 and X2; that is to see if the randomization works.

1. Let Z = 1 if a subject belongs to arm #6; = 0 if arm #5. Using X1 and X2 as covariates, fit the logistic regression to Pr(Z=1) to obtain a propensity p for each subject in both arms.
2. Let Y = log[p/(1-p)], compare the average Y-values between the two arms and draw your conclusion.
3. Use the two samples for Y to compare the two arms using the “effect size” to measure the similarity of the two arms; the effect size is defined as the difference of sample means divided by the common or average standard deviation. Would you prefer to use a test of significance, such as the two-sample t-test in (b), or a measure of difference such as the effect size to decide; explain your choice

**Problem #2:**

A phase 3, multi-center, randomized trial was conducted to compare transplantation of peripheral blood stem cells versus bone marrow from unrelated donor (as treatments for leukemia; File: BMT-Trial-E-22). Data consist of: Treatment (trttrue\_num; 1=BM, 2=PBSC), Survival time (surv; in days), Survival status (stat; 1=alive, 2-dead), and the following covariates: Gender (1=male, 2=female), Age (years), Race (white; 1= white, 0= non-white), HLA-matching (hlascore\_num; 5 or 6 out of 6).

1. Fit the Proportional Hazards Model with only one covariate, the Treatment, to obtain the Hazard Ratio (or Relative Risk).
2. Fit the Proportional Hazards Model with all covariates, including the Treatment. Which covariates are significant at the 5% level?
3. Compare the values of the Hazards Ratio (representing the Treatment Effect) obtained in (1) and (2); if they are not the same, provide your explanation.