Question 1:

Methods:

We assumed that the mean and SE provided were from a normal distribution. This allowed us to compute the Z-score (based on a normal distribution) and the 95% confidence interval (CI). We used a type I error of 5%. Importance weights for each group was determined from the proportion of Caucasian (or non-Caucasian) in each group. Efficiency weights were calculated using the inverse of the squared of the SE for each group.

(Please use this for all sub-questions in Question 1).

(a) Yes.

difference in mean cholesterol in Caucasians: -25.3 mg/dL

se = 1.5521

95% CI: (-28.34, -22.26) mg/dL

Z-score: -8.317

P<0.0001

Mean cholesterol levels were found to be 25.3 mg/dL lower in Caucasians who are male than in Caucasians who are female. This observation or it extreme was enough to rule out the null hypothesis that there was no difference in mean cholesterol level across sex for Caucasians (two-tailed, P<0.0001). From the 95% CI, the observed difference in mean cholesterol is not unusual if the true difference was anywhere between 28.34 mg/dL and 22.26 mg/dL lower in Causasian who were male than Caucasian who were female.

(b) Yes.

se = 3.453

difference in mean cholesterol in non-Caucasians: -15.7 mg/dL

95% CI: (-22.47, -8.93) mg/dL

Z-score: -4.547

P<0.0001

Mean cholesterol levels were found to be 15.7 mg/dL lower in non-Caucasians who are male than in non-Caucasians who are female. This observation or it extreme was enough to rule out the null hypothesis that there was no difference in mean cholesterol level across sex for non-Caucasians (two-tailed, P<0.0001). From the 95% CI, the observed difference in mean cholesterol is not unusual if the true difference was anywhere between 22.47 mg/dL and 8.93 mg/dL lower in non-Causasian who were male than non-Caucasian who were female.

(c) Yes.

mean difference in cholesterol in Caucasians: -25.3 mg/dL

mean difference in cholesterol in non-Caucasians: -15.7 mg/dL

Important weights:

* Wc = 0.8405
* Wn=0.1595

Adjusted difference in means between Caucasians and non-Caucasians: -23.77 mg/dL

se(adjusted): 1.416

Z-score: -16.79

P<0.0001

95%CI: (-26.55, -21.50) mg/dL

After adjustment for sex and race, mean cholesterol levels were found to be -23.77 mg/dL lower in male patients than in patients of the same race who are female. This difference was found to be significantly difference across sex (two-tailed, P<0.0001). From the 95% CI, the observed difference was not unusual if the true difference in mean cholesterol between sex was anywhere between 26.55 mg/dL lower and 21/50 mg/dL lower in males than females of the same race.

Efficiency weights: 1/(SE^2)

I get the exact same answer using efficiency weights.

(d) Yes.

difference of the difference: -9.6 mg/dL

se(effect modification): 3.786

Z-score: -2.536

P= 0. 0112

95% CI: (-11.04, -2.179) mg/dL

The difference in mean cholesterol across groups defined by race was found to be 9.6 mg/dL lower in males than in females. Such difference was sufficiently extreme to be able to rule out with high confidence a null hypothesis of no effect modification by race in the association between male and female (two-tailed, P<0.0112). Based on a 95% CI, we find that the observed difference in the association between cholesterol and sex across the race groups not atypical of settings in which the true difference in effect were such that males had mean difference in cholesterol across race groups 11.04 mg/dL lower to 2.179 mg/dL lower than that in females.

Question 2:

Methods:

We assumed that the mean and SE provided were from a normal distribution. This allowed us to compute the Z-score (based on a normal distribution) and the 95% confidence interval (CI). We used a type I error of 5%. Importance weights for each group was determined from the proportion of Caucasian (or non-Caucasian) in each group. Efficiency weights were calculated using the inverse of the squared of the SE for each group.

(Please use this for all sub-questions in Question 2).

(a) No.

Mean difference in fibrinogen in Caucasians: -2.9 mg/dL

se: 2.677

Z-score: -1.083

P=0.2788

95% CI: (-8.147, 2.347) mg/dL

Mean fibrinogen level was 2.9 mg/dL lower in Caucasians who were males than Caucasians who were females. This difference was not extreme enough to reject the null hypothesis that the difference across sex for Caucasian was not significantly difference from 0 (two-tailed, P=0.2788). From the 95% CI, the observed difference was not unusual if the true difference in fibrinogen levels across sex for Caucasians was anywhere from 8.147 mg/dL lower to 2.347 mg/dL higher in males than females who were Caucasian.

(b) Yes.

Mean difference in fibrinogen in non-Caucasian: -15.7 mg/dL

se: 7.2964

Z-score: -2.1518

P=0.03141

95% CI: (-30.01, -1.399) mg/dL

Mean fibrinogen level was 15.7 mg/dL lower for non-Caucasian males than non-Caucasian females. This difference was extreme enough to reject the null that there was no difference in fibrinogen levels across sex for non-Caucasian (two-tailed, p=0.0314). From the 95% CI, the observed difference was no unusual if the true difference in fibrinogen level across sex for non-Caucasian was anywhere from 30.01 mg/dL lower and 1.399 mg/dL lower for non-Caucasian males than for non-Caucasian females.

(c) No.

Adjusted mean fibrinogen difference: -4.9416 mg/dL

Importance weights:

* Wc: 0.8405
* Wn: 0.1595

se(adjusted): 2.533

Z-score: -1.951

P-value: 0.05106

95% CI: (-9.906, 0.02308) mg/dL

After adjustment for sex and race, the mean fibrinogen levels were found to be 4.942 mg/dL lower for males than females with the same race. This difference was found to be not sufficiently extreme enough to reject the null hypothesis that there was no difference in fibrinogen across sexes with the same race (two-tailed P=0.05106). From the 95% CI, the observed difference was not unusual if the true difference in fibrinogen level across sex for the same race was anywhere between 9.906 mg/dL lower to 0.02308 mg/dL higher in males than females of the same sex.

Efficiency weights:

W = 1/(SE^2)

Therefore, the adjusted mean difference is -4.418 mg/dL

SE(adjusted) = 2.516

Z=score: -1.756

P= 0.07909

95% CI: (-9.349. 0.5134) mg/dL

After adjustment for sex and race, the mean fibrinogen levels were found to be 4.418 mg/dL lower for males than females with the same race. This difference was found to be not sufficiently extreme enough to reject the null hypothesis that there was no difference in fibrinogen across sexes with the same race (two-tailed P=0.07909). From the 95% CI, the observed difference was not unusual if the true difference in fibrinogen level across sex for the same race was anywhere between 9.349 mg/dL lower to 0.5134 mg/dL higher in males than females of the same sex.

(d)

difference of the difference: 12.8 mg/dL

se: 7.772

Z-score: 1.647

P-value: 0.9958

95% CI: (-2.433, 28.04) mg/dL

The difference in mean fibrinogen across groups defined by race was found to be 15.7 mg/dL lower in males than in females. Such difference was not sufficiently extreme to be able to rule out with high confidence a null hypothesis of no effect modification by race in the association between male and female (two-tailed, P<0.9958). Based on a 95% CI, we find that the observed difference in the association between fibrinogen and sex across the race groups not atypical of settings in which the true difference in effect were such that males had mean difference in fibrinogen across race groups 2.433 mg/dL lower to 28.04 mg/dL higher than that in females.

Question 3:

Methods:

Descriptive analysis evaluated the number, mean, sd, and range of the cholesterol levels in the sample. A one-sided level of significant of 0.025 was used for sample size calculations. A power of 80% or 90% was used to estimate sample size calculation where appropriate (as determined by the sub-question). Therefore, Z(1-alpha) = 1.96, Z(beta=80%) was 0.8416, and Z(beta=90%) was 1.282.

The proportion of patients who had less than 200 mg/dL cholesterol was defined as the expected proportion of control patients and was a binary variable. The proportion of patients who had less than 210 mg/dL cholesterol was defined as the expected treatment group and was a binary variable.

(Please use this method for all subgroup questions).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 1. Descriptive analysis of cholesterol in the sample (N=5000). | | | | | | |
|  | N | Mean | SD | Min | Max | Missing |
| Age (years) | 5000 | 72.83 | 5.596 | 65 | 100 | 0 |
| Cholesterol (mg/dL) | 4953 | 211.69 | 39.29 | 73 | 430 | 7 (0.14%) |
| Fibrinogen (mg/dL) | 4915 | 322.98 | 67.29 | 109 | 872 | 85 (1.7%) |
| BMI (m/kg^2 | 4987 | 26.67 | 4.735 | 14.7 | 58.8 | 13 (0.26%) |
| Systolic BP (mm Hg) | 4990 | 136.55 | 21.86 | 77 | 235 | 10 (0.2%) |
|  | N | % |
| Male | 2096 | 41.92 |
| Female | 2904 | 58.08 |
| Smoker (Yes) | 604 | 12.08 |
| No | 4390 | 87.8 |
| Missing | 6 | 0.12 |

There was 7 missing data for cholesterol, which accounted for 0.002% of the population. A total of 4953 (99.06%) of the population had complete cholesterol data. For fibrinogen there was 85 (1.7%) missing data. For BMI, there was 13 (0.26%) missing data. For systolic BP, there was 10(0.2%) missing data. For smoker status, there was 6 (0.12%) missing data.

The average patient had a mean cholesterol of 211.69 mg/dL, mean fibrinogen of 322.98 mg/dL, mean BMI of 26.67 mg/dL, and systolic BP of 136.55 mg/dL. A majority of the sample was female (58.08%) and non-smoking (87.8%).

(Please use the above Methods for Questions 3 to 5).

(a) Best estimate of the SD of cholesterol within sample: 39.29 mg/dL

(b) SD of the change in cholesterol measurement made after 3 years within the population: 43.04 mg/dL

V = sigma1^2 + sigma2^2 – 2(rho)(sigma1)(sigma2)

Assuming equal variances and sigma = 39.29 mg/dL

SQRT(V) = SD = 43.04 mg/dL

(c) Best estimate of the SD of cholesterol within groups that had constant age and sex: 37.49 mg/dL or RMSE

Question 4:

(a) Sample size with 80% power to detect the alternative: 582

(b) Sample size with 90% power to detect the alternative: 779

(c) Sample size would decrease from 582 to 530 if age and sex were not adjusted.

(d) Since we are not considering the within group change, we lost precision in our estimate. Analyzing only the final cholesterol between the two groups does not take the baseline into account; hence losing precision. Consequently, we need fewer patients to maintain 90% power.

(e) If the correlation is less than 0.5, ANOVA won’t be affected by baseline cholesterol in an RCT. Since we have an RCT, we should not expect much difference with the results from an ANCOVA.

Question 5:

(a) Estimate of the proportion of subjects on the control arm with cholesterol less than 200 mg/dL: 0.3920

(b) Estimate of the proportion of subjects on the treatment arm with cholesterol less than 210 mg/dL: 0.4896

(c) Sample size with 90% power to detect the alternative: 949

(d) For both the distribution free inference is based on the CLT.

Effect modification: The disadvantages of using this study design are that the precision is decreased when the correlation is 0.5. We also need to know the mean-variance relationship and the relative good size. We need to know both proportions for the two groups in order to know V.

Adjustment: In the previous method, we perform a randomized study. In the previous method, we would need to know the variance of observations within groups, variance of predictor X, and correlation of X on any additional covariates. Precision is related to within group variance (when adjusted, increases precision), and the range we measure, and the correlation with the POI and variables used in the adjustment (we may not always have this information), which can lead to decreased precision. Interpretability may be easier with the previous method.