**Biostatistics 576B**

**Homework 5**

The following data are from a case control study of the relationship of recent oral contraceptive use and myocardial infarction.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Recent  OC  Use |  | |  | | Age Category | |  | |  | |
| 25 - 29 | | 30 - 34 | | 35 - 39 | | 40 - 44 | | 45 - 49 | |
| Case | Control | Case | Control | Case | Control | Case | Control | Case | Control |
| Yes | 4 | 62 | 9 | 33 | 4 | 26 | 6 | 9 | 6 | 4 |
| No | 2 | 224 | 12 | 390 | 33 | 330 | 65 | 362 | 93 | 301 |

Analyze the data using a stratified analysis.

1. Determine whether the odds ratios are homogeneous across the age strata.

kable(res$df1, align = res$df1.align)

P.estimate |Stats |95%CI-ll |95%CI-ul |

|:--------------------------------|:------|:--------|:--------|

|MH test of Homogeneity (p-value) |0.22 | | |

|Crude OR for OC |1.70 |1.07 |2.63 |

|MH OR OC adjusted for age |4.13 |2.52 |6.78 |

|Adjusted/crude relative change |143.38 | |

Since our MH test for homogeneity is .22, this pvalue is small enough that we would still report each stratum. However, we there is not enough evidence to conclude the odds ratios are different across stratum.

1. If appropriate, determine whether age is a confounder of the oral contraceptive use versus myocardial infarction relationship.

Since the crude OR is 1.7 and the adjusted OR is 4.13, since this increase is greater than 10% age may be a confounder for OC use and myocardial infarction.

1. Test whether there is an association between oral contraceptive use and myocardial infarction (adjusting for age if necessary).

The 95% confidence interval for the adjusted odds ratio is 2.52 to 6.78 which does not include 1 thus there is an association.

Now use logistic regression to perform the same analysis.

1. Determine whether the odds ratios are homogeneous across the age strata.

Call:

glm(formula = disease ~ OC + age + iter, family = binomial(link = "logit"),

data = df\_1)

Deviance Residuals:

Min 1Q Median 3Q Max

-1.2576 -0.5503 -0.3911 -0.2754 2.8260

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -4.69565 0.30149 -15.575 <2e-16 \*\*\*

OC 1.50519 0.59223 2.542 0.011 \*

age 0.72116 0.07199 10.018 <2e-16 \*\*\*

iter -0.04575 0.18331 -0.250 0.803

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1437.4 on 1974 degrees of freedom

Residual deviance: 1286.7 on 1971 degrees of freedom

AIC: 1294.7

Number of Fisher Scoring iterations: 6

Since our interaction term is not significant, we can not say the odds ratios are different across age strata.

1. If appropriate, determine whether age is a confounder of the oral contraceptive use versus myocardial infarction relationship.

> or\_glm(df\_1, OR\_log, incr = list(OC=1))

# A tibble: 5 x 5

predictor oddsratio `CI\_low (2.5)` `CI\_high (97.5)` increment

*<chr>* *<dbl>* *<dbl>* *<dbl>* *<chr>*

1 OC 4.13 2.50 6.73 1

2 as.factor(age)2 3.15 1.31 8.79 Indicator variable

3 as.factor(age)3 7.00 3.06 19.0 Indicator variable

4 as.factor(age)4 14.3 6.41 38.6 Indicator variable

5 as.factor(age)5 25.0 11.2 67.0 Indicator variable

Thus our adjusted odds ratio is 4.13 and our crude is 1.7, thus age is likely a confounder.

1. Test whether there is an association between oral contraceptive use and myocardial infarction (adjusting for age if necessary).

Adjusting for age, the OC predictor has a significant pvalue, thus we can say there is an association between oral contraceptive and myocardial infarction.

summary(OR\_log <-glm(disease~OC+as.factor(age), data=df\_1, family = binomial(link= "logit")))

Call:

glm(formula = disease ~ OC + as.factor(age), family = binomial(link = "logit"),

data = df\_1)

Deviance Residuals:

Min 1Q Median 3Q Max

-1.2852 -0.5729 -0.4085 -0.2772 2.9665

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -4.3876 0.4357 -10.071 < 2e-16 \*\*\*

OC 1.4191 0.2516 5.640 1.70e-08 \*\*\*

as.factor(age)2 1.1475 0.4773 2.404 0.0162 \*

as.factor(age)3 1.9461 0.4588 4.241 2.22e-05 \*\*\*

as.factor(age)4 2.6637 0.4505 5.913 3.36e-09 \*\*\*

as.factor(age)5 3.2184 0.4485 7.175 7.22e-13 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1437.4 on 1974 degrees of freedom

Residual deviance: 1284.5 on 1969 degrees of freedom

AIC: 1296.5

Number of Fisher Scoring iterations: 6

1. Compare your results using both approaches.

Both results gave us the same conclusion and both methods gave us an adjusted odds ratio of 4.13!