In question 5, our task is to examine the role of gender in the obesity-depression causal relationship. Consider that for each gender we can generate a 2×2 cross tabulation to show the number of people with different health conditions, instead of comparing the results of Fisher exact test of each cross tabulation, we used Cochran-Mantel-Haenszel Chi-Squared Test (CMH Test) to test whether gender has a statistically significant impact on the relationship between obesity and depression.

*Table 1: Cross tabulation of gender, obesity and depression*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Obesity** | **Depression** | **Male** | **Female** |  |
| Yes | Yes  No  Percentage | 491  1189  29.2% | 875  1257  41.0% |  |
| No | Yes  No  Percentage | 779  3236  19.4% | 1260  3357  27.2% |  |

After removing rows containing meaningless values and transforming column “HLT\_BMI\_CAT” into a binary variable, we have three nominal variables: gender (male or female), obesity (yes or no) and depression (yes or no). The table above shows the number of persons with different diseases. There is a smaller proportion of depression rate in male; we want to know whether this difference is significant.

We set the null hypothesis that the proportion of people with depression is the same in male and female (true common odds ratio is equal to 1), and got a Mantel-Haenszel chi-squared of 192.26 with p-value less than 2.2e-16. Thus, we made the statement that the differences caused by gender are statistically significant.