

Example Summary of This Final Analysis

To examine the relationship between smoking status (0 = nonsmoker, 1 = former smoker, 2 = occasional smoker, 3 = daily smoker) and the age of onset of diabetes (DIBAGE) for participants aged >20 years, smoking status was converted to three weighted effects coded variables (with nonsmokers as the comparison group). Each effects coded variable (WEFF1–WEFF3) was weighted for the comparison group by the ratio of the size of each other group ($N = 818$, 92, and 337, respectively) to that of the comparison group ($N = 1,206$). This produces comparisons between each of the other groups (former smokers, occasional smokers, and daily smokers) and the estimated population mean.

Assumptions of OLS regression were checked and were reasonably met. Specifically, the variances of the different groups were reasonably close, and the residuals were normally distributed (skew = -0.036 , kurtosis = -0.47).

Overall, the model was significant, and smoking status accounted for about 4.7% of the variance in age of onset ($R = 0.216$, $R^2 = 0.047$, $F_{(3, 2,449)} = 40.04$, $p < .0001$). Each comparison, as shown in Table 4.13, was significant, indicating that each of the latter three groups (former smokers mean = 54.52, occasional smokers mean = 46.91, daily