# Part 1 – Introduction & Preliminary analysis

Data on 2,700 individuals were collected through an observational study. Information on age, sex, education level, smoking status, and waist to hip ratio was recorded for all 2,700 individuals. There are two outcomes of interest – a binary carotid plaque status (plaque absent or present), and a whole-number valued plaque count. Plaque count was recorded for 30% of the sample. Table 1 shows descriptive summary statistics for these variables.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Carotid plaque status (N=2700)* | | *Carotid plaque count (N=810)* | |
|  | N | Plaque present, % (N) | N | Plaque count, mean (SD) |
| Age 40-49 | 803 | 42.2% (339) | 245 | 0.7 (0.9) |
| Age 50-59 | 863 | 60.8% (525) | 256 | 1.3 (1.5) |
| Age 60-69 | 1,034 | 79.8% (825) | 309 | 2.1 (1.7) |
| Sex Female | 1,550 | 55.4% (858) | 469 | 1.0 (1.2) |
| Sex Male | 1,150 | 72.3% (831) | 341 | 2.0 (1.7) |
| Edu. >= Univ. | 1,725 | 66.2% (1,142) | 510 | 1.6 (1.6) |
| Edu. < Univ. | 975 | 56.1% (547) | 300 | 1.2 (1.4) |
| Smk. Never | 1,392 | 57.6% (802) | 424 | 1.1 (1.3) |
| Smk. Ex | 661 | 64.0% (423) | 202 | 1.5 (1.6) |
| Smk. Current | 647 | 71.7% (464) | 184 | 2.0 (1.8) |

presence of carotid plaque increases with age, is more frequent in males compared to females, and more frequent in current & ex-smokers compared to never smokers. Additional cross-tabulations (not shown here) show that there are interactions between these variables. In this sample males are significantly more likely to smoke than females, the proportion of university educated individuals is significantly lower in older age groups, and a significantly higher proportion of the older age groups smoke compared to younger groups. There is overdispersion in plaque count, which is addressed in part 4.

# Part 2 – Smoking & presence of carotid plaque

This section examines the relation between smoking and presence of carotid plaque. A logistic regression model is used to adjust for confounding due to age, sex, and education (a proxy for social class). First the data is aggregated by the confounders, and initial models are fit on the grouped data. Deviance is used to assess model fit, and more complex forms of the linear predictor are assessed through likelihood ratio tests. The model was than fit to the individual level data and used to estimate marginal odds ratios of carotid plaque presence in ex-smokers and current smokers, compared to never smokers. Adjusted marginal probabilities for presence of plaque were calculated using the margins command and by empirical standardisation. Finally, the model was used to estimate prevalence of carotid plaque presence if everyone in the sample never smoked.

An initial binomial model was fit to the grouped data. The model uses the individuals smoking status, age, sex, and level of education (all encoded as categorical variables), with no interactions between any of the variables. The deviance of this model is 63.6 on 62 degrees of freedom, comparing this to a distribution gives a p-value of 0.42 – there is no evidence of poor model fit.  
More complex models were also considered. Three interaction terms were separately added to the initial model – an interaction between smoking and age, between age and education, and between smoking and sex. A fourth model with interactions between all variables was also tried. None of the models had significantly better fit relative to the initial no-interaction model, so the additional terms were not included (p-values from likelihood ratio tests were 0.66, 0.89, 0.16, and 0.42 respectively).

The initial model was fit to the individual data and used to produce marginal probabilities of plaque presence. Marginal odds ratios are also calculated and are shown in table 2.

|  |  |  |
| --- | --- | --- |
| Smoking status | Marginal probability | Marginal odds ratio |
| Never smoker | 0.59 [0.57, 0.62] |  |
| Ex-smoker | 0.61 [0.58, 0.65] | 1.12 [0.90, 1.33] |
| Current smoker | 0.72 [0.68, 0.75] | 1.92 [1.43, 2.17] |
| **Table 2**: Estimated marginal probability of plaque presence by smoking status. Notes: 95% confidence intervals shown in square brackets. Estimates are adjusted for age, sex, and education. Estimates are rounded and may not line up with calculations quoted in the text. Odds ratio CIs based on percentile bootstrap estimates. | | |

The marginal probabilities in table 2 can be interpreted as prevalence of carotid plaque in the situation where all study participants have the same smoking status. For example, the prevalence would be 59% if no one in the study smoked. Assuming smoking has a causal effect on carotid plaque, this means that 3.5% (95% CI [1.0%, 6.0%]) of carotid plaque prevalence is due to smoking. If no one in the study ever smoked, and assuming that the effect of smoking is causal, there would be between 27 and 163 fewer carotid plaque cases.

# Part 3 – Waist to hip ratio & presence of carotid plaque

Next, waist to hip ratio (W2H) is added to the model from part 2, and the effect of adding W2H is quantified using a likelihood ratio test. The need for potential interactions were examined through lowess plots. Interactions between W2H and smoking, and W2H and age were examined using likelihood ratio tests but not added. This ‘no-interaction’ model is used to quantify the effect of increasing W2H has on presence of carotid plaque.  
An interaction term between W2H and sex is then added, in order to examine how the association between W2H and presence of carotid plaque varies by sex. This model is then used to examine the prevalence of carotid plaque if the study population all followed WHO guidelines on obesity.

A likelihood ratio test between the models with & without W2H was highly significant (19.9 on 1 degree of freedom, p < 0.0001), there is strong evidence that adding W2H improved model fit. There is no evidence that model fit is improved by including interaction terms between W2H and smoking, or W2H and age (likelihood ratio test p-values 0.42 and 0.83 respectively). As such, interaction terms were not added to the ‘no-interaction’ model.

The no-interaction model estimates that an increase in W2H of 0.1 leads to an increase in the odds ratio of having carotid plaque by 1.30 (95% CI [1.16, 1.46]), adjusting for age, education level, and smoking status. Including an interaction between sex and W2H does not significantly improve model fit (LR test p-value = 0.69) but is included to quantify how the effect of W2H varies by sex. This model estimates that a 0.1 unit increase in W2H leads to an increase in the odds ratio of having carotid plaque by 1.28 (95% CI [1.11, 1.47]) in females, and by 1.34 (95% CI [1.11, 1.62]) in males. These estimated odds ratios are adjusted for the same variables as the no-interaction model. There is no evidence that the effect of W2H varies by sex (coefficient Wald test p-value = 0.69).

The effect of reducing levels of obesity was quantified. Values of W2H which are classed by WHO as obese were reduced – males with W2H above 0.9 had their W2H value set to 0.9, and females with W2H above 0.85 had their W2H values set to 0.85. The ‘sex-interaction’ model was used, as the WHO guidance varies by sex. If no one in the study was obese, the prevalence of carotid plaque presence is estimated at 60.6% (95% CI [58.6%, 62.5%]).

# Part 4 – Smoking and waist to hip ratio & plaque count

This section looks at the 810 study participants (30%) with plaque count recorded. Three negative binomial models of plaque count are fit – (1) age, sex, education, and smoking, (2) age, sex, education, smoking, and W2H, and (3) model 2 with an interaction between W2H & sex. Negative binomial models are used due to the overdispersion noted in part 1. Interactions between age & sex, and age & smoking status were tested using likelihood ratio tests, but not added as neither interaction led to a significant improvement in model fit.

Table 3 shows the coefficient estimates, 95% confidence intervals, and p-values from the three models. These estimates are the change from baseline in mean plaque count.

# Part 5 – Discussion and conclusions

A series of models have been used in this report to estimate the causal effect of smoking and waist to height ratio on carotid plaque. Table X summarises the effect estimates:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| (a) Effect on carotid plaque prevalence | | | | | |
| **Model** | **Variable** | **Odds ratio** | **95% CI** | **p-value** | **Adjusted for** |
| 1 | Ex-smoker | 1.12 | [0.89, 1.40] | 0.34 | Age, sex, education |
| Current smoker | 1.92 | [1.52, 2.43] | <0.0001 |
| 2 | W2H | 1.30 | [1.16, 1.46] | <0.0001 | Age, sex, education, smoking |
| 3 | W2H – females | 1.28 | [1.11, 1.47] | 0.001 |
| W2H – males | 1.34 | [1.11, 1.62] | 0.003 |
| (b) Effect on carotid plaque count | | | | | |
| **Model** | **Variable** | **Δ plaque** | **95% CI** | **p-value** | **Adjusted for** |
| 4 | Ex-smoker | 0.47 | [-0.12, 0.22] | 0.59 | Age, sex, education |
| Current smoker | 1.50 | [1.27, 1.78] | <0.0001 |
| 5 | W2H | 1.11 | [1.02, 1.22] | 0.02 | Age, sex, education, smoking |
| 6 | W2H – females | 1.22 | [1.08, 1.37] | 0.001 |
| W2H – males | 1.01 | [0.89, 1.15] | 0.84 |
| **Table X**: Effect estimates. Notes: Δ plaque is the change from baseline in mean plaque count. Effect estimates in models 1 and 4 are relative to never-smokers. Effect estimates in models 2,3,5,6 are for an increase of 0.1 in W2H. All effect estimates in the table are conditional on the adjustment variables. Coefficients for adjustment variables not shown as they do not have a causal interpretation. Models 3 and 6 have an interaction term between W2H and sex, all other models are linear in the variables. Model 1 fit on grouped data, all others on ungrouped data. | | | | | |

Being a current smoker causes the largest increase on carotid plaque count and the odds of plaque presence. The prevalence of carotid plaque was estimated in two scenarios – one where the study population were all never-smokers, and one where no-one in the study population is obese. Reducing smoking levels led to a larger reduction in plaque prevalence compared to reducing obesity levels. This suggests that smoking cessation services would be more effective at reducing plaque prevalence compared to weight management services, which could help local public health teams to prioritise interventions aimed at reducing carotid plaque prevalence in the population.

The data in this study is observational, this adds some limitations to our results. Throughout the report, the effects in table X have been interpreted causally. This relies on two strong assumptions. The first is that there is no unmeasured confounding – there are no additional confounding variables which haven’t been adjusted for in the models. Level of education is a proxy of social class and all models in this report adjust for education, aiming to remove any confounding due to social class. If there are other social class confounders which are not causally related to education level, then they will not be properly adjusted for through the education variable. This will introduce bias in our results.  
The second assumption is that the causal relationships between the variables have been correctly captured. A causal diagram / DAG was not included in the assessment, so this point can’t be critiqued in detail. Figure 1 shows one potential DAG

Diagram

Description automatically generated

This captures the associations mentioned in part 1. If this is the true DAG then a model adjusting for age and smoking, or sex and smoking would correctly identify the causal effect of W2H on carotid plaque. To estimate the causal effect of smoking the only adjustments needed would be age and sex. There is a chance that the wrong variables were adjusted for in the W2H estimates. This is another potential source of bias.

An RCT would be unethical due to the nature of the intervention, so these biases can never be eliminated through study design. The strength of evidence could be improved through other observational studies – for example are the effects seen in this study replicated in other similar studies on different populations?