Periodontal Disease and CHD Analysis and Report Assignment

An association between periodontal disease and cardiovascular disease

(CVD) has been observed in some large, observational cohort studies.

Some researchers have postulated that the inflammatory process that

occurs with periodontal disease has a damaging effect on blood vessels

and the heart via inflammatory cytokines (biologically active

cytokines that participate in the body's defense against foreign

invasion and might cause "collateral" damage to the body itself while

fighting off pathogens.) C-reactive protein (CRP), for example, has

been found to be associated with elevated risk of hearth attack in a

dose-response manner.

You have available to you data from a large, observational study that

includes information on hospitalization and death due to coronary

heart disease (CHD), as well as information on periodontal disease

status. \*\* Your assignment is to use these data to assess the strength

of evidence for a causal association between periodontal disease and

CHD. Choose the statistical methods you see fit for this purpose,

given the data available.\*\* Write a report of your findings in the

style of: Abstract, Introduction, Methods, Results, Discussion,

Conclusion. Use language, tables and figures that will be clear to a

college graduate with a health sciences background. You can assume

the audience has some background on CHD, so don't write a report on

"What is CHD," though you may find that a bit of personal background

reading to be helpful.

Data set: perio-data.txt.

Documentation on variable definitions: perio-doc.txt

Some background on the data: The data were collected as part of a

large, longitudinal survey study of US residents. Subjects were

selected for the study using a cluster-sampling design, but you do not

have information on the sampling scheme for this assignment. Also, we

are interested only in baseline variables as potential predictors of

outcome (i.e. this data set does not include longitudinal data).

Hence, all variables except the outcome variables were assessed at

baseline. The exception is data on smoking: data on smoking were

collected approximately 10 years after the baseline assessment, but

subjects were asked to recall their smoking behavior at baseline.

Hence, the smoking variables are measures of smoking exposure at

baseline, but they are subject to additional error relative to

variables actually assessed at baseline. (You might wish to discuss the

ramifications of measurement error in your report.)