# Illustrative Example of Study Protocol  - Matthew Staley - May 1999

1.      Research Question and Hypothesis

Liver cancer is a rare disease in the United States, having an incidence of only 5 per 100,000. However, despite its rarity, it is an especially severe form of malignancy having an interval from diagnosis to death of between four and six months.  Even with the most radical of treatment options (massive chemotherapy, radiation and transplant) survival rarely increases beyond eighteen months.  The relationship between chronic Hepatitis B and liver cancer has been clearly established, as have relationships with age (60-70 years), male sex and cirrhosis.  The relationship between Hepatitis C (AKA Non A, Non B hepatitis) is less clear.  It is speculated that there may be a relationship between exposure to Hepatitis C virus (HCV) and liver cancer, but this has not been conclusively shown to be true.  This study proposes to examine the relationship between HCV and liver cancer.  The null hypothesis is that there is not a significant relationship.  The alternative hypothesis is that there is a significant relationship between HCV and subsequent liver carcinoma.

2.      Target Population

The population to be sampled will include individuals diagnosed with liver cancer between 1992 and 1998 as reported to the National Cancer Registry.  This population must have been born and resided for the entirety of their lives in the United States.  Subjects will be recruited from the NCR database files and will not be controlled for age, race, SES or any other factors beyond a diagnosis of liver cancer.  Subjects need not be living, but it is necessary for them to have medical records, and preferably, serum or tissue samples available for further analysis as needed.  Controls will be matched for age, race, SES, presence/absence of cirrhosis, HBV exposure and sex.  Controls will be recruited by random sampling of the matching population using Census bureau and medical record information.

3.      Independent (Explanatory) Variable

The exposure variable will be positive HCV serostatus as determined by EIA analysis of serum samples using the FDA licensed Abbott Laboratories kit in HCFA certified laboratories.  If there are sufficient number of serum, or tissue samples available, HCV status will be confirmed by utilizing Roche Diagnostics polymerase chain reaction (PCR) based techniques for the detection of HCV RNA.  The EIA method is capable of detecting only total antibody to HCV.  It does not distinguish between IgG and IgM which would allow for a determination of recent versus past exposure, nor does antibody testing allow for an accurate assessment of the chronicity of the infection.  PCR analysis, however, will only be positive when there is actual viral RNA present in either serum or tissue samples which is a more reliable indicator of current, or chronic infection.  PCR analysis also has the advantage of being reliable when applied to paraffin tissue blocks that have been formerly subjected to histologic analysis.  It would be expected that cancer patients would have such tissue samples available.  Despite the different degrees of sensitivity and specificity associated with these two different diagnostic techniques, positive results will be considered equal indicators of exposure.

4.      Dependent (Outcome) Variable

The outcome variable is a diagnosis of liver cancer.  Liver cancer diagnosis is based on recorded information available from medical records and from the NCR.  Diagnosis must have been made by a physician or pathologist based on defined criteria accepted in the oncology community.

5.      Extraneous Variables (Potential Confounders)

Hepatitis B serostatus of cases and controls will be collected from medical records and laboratory files.  The sex of participants, presence or absence of cirrhosis of the liver, racial background, geographic area of residence, SES and age at time of diagnosis with liver cancer will all be collected.

6.      Study Design

        This is an observational study.  Its directionality is case-control, timing is historical and subject selection is fixed-disease.

7.      Planned Analyses

        Data will be described by frequency of exposure among cases and controls.  As the study will yield binary data from two independent groups from a case-control design, odds ratios and 95% confidence intervals will be calculated.  Stratification by several different variables will be essential.  Of particular interest is the possibility of interaction between HBV and HCV in promoting liver cancer.  Other possible interaction variables include age and sex, therefore stratification by those variables will be essential as well.  The presence or absence of interaction will be determined by the calculation of Chi-Square tests for interaction.  Comparison of crude and adjusted odds ratios will provide insight as to the presence or absence of confounding.

8.      Sample Size Considerations

        Sample size requirements were determined using the Epi Table feature in the Epi Info 6 analysis program available from the Centers for Disease Control and Prevention.  Because liver cancer is a rare disease, the ratio of controls to cases was set at 2.0.  Hepatitis C is also a rare exposure, occurring in less than 1% of the population of the United States.  However, HCV is not commonly tested for, although this is changing, and not all areas of the United States have access to the best techniques for detecting the presence of the virus.  In approximately 50% of cases, HCV infection can be expected to resolve with no immediately obvious consequences and diagnosis may therefore be missed.  Given these factors, percent exposure among controls was set at 1%.  Calculations of sample size requirements were pursued for several different odds ratios.  Although an odds ratio as low as 1.3 would be ideal for a disease as severe as liver cancer, the rarity of both disease and exposure would require a sample size of over 1 million which is clearly impractical, if not impossible given the geographic constraints.  Raising the odds ratio to 3.0 requires 69,768 cases and 139,536 controls.  This would be possible if historical data from the time frame indicated earlier could be collected.  However, given the rarity of disease and exposure, the time, effort and financial resources required to conduct a study of this enormity, a pilot study seems essential.  If a significant association between HCV and liver cancer were indicated by the pilot, a study of larger scope would be justified.

9.      Study Limitations

        Information bias is a concern given that two different methods of determining exposure are allowed.  Choosing only one method could eliminate this, but the more widely available method is not as accurate, while the more accurate method is restricted in its availability and is prohibitively expensive at approximately $150.00 per analysis.  Selection bias is a concern because those cases of liver cancer detected and reported to the NCR may be biased as to SES, race and geographic location.  Controls may be biased due to the fact that participation is voluntary and also due to the need for historical records of exposure to HCV which is not commonly screened for in the absence of some diagnostic suspicion.  Eliminating bias from case selection seems unlikely given the need for a large number of cases of a rare disease which are only likely to be available readily from a source such as the NCR.  Controls could be selected on a non-voluntary basis if they were patients of an academic medical center where patients are required to sign a release at time of admission which gives permission for samples to be tested later without consent.  As HCV testing is not subject to consent requirements, as HIV is, this would be possible, but another form of selection bias is introduced since patients of academic medical centers tend to more ill than those utilizing other facilities.  Given these constraints, results will probably not be readily generalizable across all the population of the United States, but results would still be of interest in that an established association between HCV exposure and liver cancer might prompt more screening for HCV and closer attention to the early, non-specific symptoms of liver cancer among those who have an exposure to HCV.