3.0 DESCRIPTION AND JUSTIFICATION OF HHHQ AREAS OF INQUIRY

With few exceptions, HHHQ non-diet questions are taken directly from pretested questionnaires used in large-scale surveys or standardization projects. Page and question numbers correspond to those in Version 02.1, March, 1987 (FULL87) of the questionnaire (Section 6). Some of the questions described below are currently included only in the FULL87 version of the questionnaire.

3.1 MISCELLANEOUS INFORMATION (PAGE 1)

The following data serve a quality control or file management function:

A. Individual ID Number: Use of check digit is recommended, for file quality control. The respondent's initials may be incorporated as part of the ID number, instead of a check digit, to minimize errors.

B. Date of Interview: A time referent for future prospective uses of the data. In conjunction with question 2, an internal cross-check on date of birth.

C. Respondent: Permits description of data sources, separate examination of data from proxy respondents.

D. Coder ID: (Page 10) Permits quality control, inter-coder reliability checks.

3.2 IDENTIFIER INFORMATION (PAGE 1)

Disease prevention and control and other health policies will entail the application of intervention strategies on a very large scale (conceivably the entire population) and over a very long time period (conceivably an individual's entire lifetime). Both scientific and ethical considerations demand that we monitor the long-term effect of such interventions. Monitoring of causes of death can be achieved through matching with the National Death Index (NDI). Some adverse or beneficial health effects may not be manifested in mortality, however, and it is thus also important to facilitate tracing for long-term morbidity evaluation.

A. Name: Full name is required for matching with the NDI. For females, maiden name also facilitates tracing and is one of the NDI matching factors.

B. State of Birth: An NDI matching factor.

C. Address and Telephone Number:

1. Tracing information is required for long-term morbidity evaluation or follow-up; to facilitate location of the individual during the course of the study to maintain adverse effect and efficacy monitoring; to ensure attribution of the questionnaire to the correct individual; and to facilitate queries to the respondent regarding missing data.

2. Inclusion of county and zip code information permits comparison of results with nationally published rates of disease by geographic area, or by other parameters such as population density.

NOTE that the FULL87 questionnaire is formatted to permit the removal of the page containing identifier information, for confidentiality.

3.3 DEMOGRAPHIC DATA (PAGE 3, QUESTIONS 1-7)

Variables such as age, sex and race are basic to the description of any population and would be examined in any comparison between populations. This reflects their acknowledged importance in virtually every health outcome. Related socioeconomic variables such as education, marital status, and mobility have also been shown to be associated with health outcomes, and thus it is important to include them in descriptions of the population and consider controlling for them in the analysis, if their distribution is uneven.

A. Date of Birth: A confounder which must be controlled for in all analyses. Required for matching with the NDI.

B. Age: A potential confounder.

C. Race: A potential confounder. These are the standard categories prescribed by the Department of Health and Human Services.

D. Education: A marker of socioeconomic status, more interpretable over time than income. With Q. 27, Occupation, a modified McGuire and White (1955)/Warner et al. (1949) Index of Social Status may be constructed.

E. Marital Status: Shown to be significantly associated with health and mortality outcome, by Berkman and Syme (1979), Wingard et al. (1982), House et al. (1982), and many other investigators.

F. Mobility: Mobility was shown by Haenszel et al. (1962) to be associated with lung cancer mortality after controlling for smoking and other factors. It may be a proxy measure for stress or socioeconomic factors.

3.4 SMOKING (PAGE 3, QUESTIONS 8-9)

"Cigarette smoking is the single most important environmental factor contributing to premature mortality in the United States", according to the Surgeon General's Report on Smoking and Health. Most studies will examine total mortality and total morbidity, even where that is not the central focus of the study, and significant policy implications may be derived from such an examination. Thus, it is important to control for the potential confounding by this factor in all studies.

A. Smoking-ever: In this question, and most of the subsequent smoking questions, the wording is that used by the U.S. National Health Interview Survey (NHIS) prior to 1991. The preservation of identical definitions and wording will permit comparison of each study sample with a large annual national probability sample. It would be prudent for investigators to look at the 1993 and future U.S. NHIS and consider appropriate adjustments in these questions.

B. Age Started Smoking: This has been shown to be an important variable in health outcome. This question will permit the calculation of lifetime pack-years, which in turn will permit comparison with many studies in which this summary variable is used.

C. Current Smoking Status and Amount: This will permit comparison of study groups, comparison with rates and relative risks found in other studies, and comparison with rates found in the National Health Interview Survey. (See A above about newer NHIS definitions since 1991.)

D. Average Amount Smoked (Lifetime): This represents the usual lifetime smoking exposure, and is the form recommended by the American Thoracic Society Standardization Project.

E. Age Stopped Smoking: Permits the calculation of pack-years of exposure.

3.5 DIETARY INTAKE (PAGES 3-7, QUESTIONS 10-22)

Of all the factors known to be associated with disease risk, several investigators have estimated that diet may have the single largest effect. The role of energy, fat and cholesterol intake in heart disease has long been recognized, and the role of antioxidants is being increasingly appreciated. For cancer, Doll and Peto (1981) estimate that 35 percent of all cancer deaths may be attributed to diet. In addition to studies of the role of dietary factors themselves, diet is a potential confounding factor which may influence the apparent results of an investigation of other factors. In addition, total dietary adequacy or specific dietary factors may be modifiers, altering the effectiveness of intervention agents or other exposures. Consequently, interpretations of the results of any study of disease will be enhanced by a knowledge of the distribution of this risk factor in the study populations. In addition, diet is a risk factor for other major causes of morbidity and mortality, and may affect host resistance or response to harmful or protective agents.

The dietary assessment instrument should be one which provides information about an individual's -- as opposed to a group's -- usual diet. (See Hartman and Block, 1992, for a discussion of methods and statistical issues.) These requirements, and the practical considerations for large-scale epidemiologic studies, indicate that a quantified food-frequency questionnaire is the appropriate method. See Section 4, Introduction to the Diet Questionnaire for a discussion of the dietary assessment methodology and evaluation of its validity.

3.6 HEALTH HISTORY (QUESTIONS 23-26 AND 31-34)

A history of major illness is associated with adverse health and mortality outcomes, and thus an uneven distribution of such risk factors in the study population should be examined and controlled for, if appropriate. Additionally, some factors may be indicators of a premalignant condition or may represent a marker or proxy for dietary inadequacy or metabolic defect. A knowledge of their existence and distribution in the study groups would facilitate interpretation of results of the study.

A. Hospitalizations: An objective measure of health status not covered in Q. 24-25.

B. Medical Radiation: Radiation, including explicitly some of these medical exposures, is a well-demonstrated risk factor in several cancer sites, and the number of cancer sites associated with radiation exposure is increasing as additional studies are reported. The capability of describing these pre-existing risk factors in the study groups and of controlling their potential confounding effect enhances interpretability of the results.

C. Disease Checklist: As indicated above, disease history is a predictor of subsequent morbidity and mortality. Q. 25 is a slightly expanded form of the list used by the American Cancer Society's nationwide Cancer Prevention Study II.

D. Height-Weight: (Page 10) Obesity or factors associated with it have been implicated in a number of cancers, as well as other major diseases, and weight status was one of the five health practices found to have a significant independent association with subsequent mortality in the Alameda County study discussed further below.

3.7 OCCUPATIONAL DATA (PAGE 9, QUESTIONS 27-29)

Because of the potential carcinogenicity of occupational exposures, it is important to evaluate their importance in any study in which cancer or total mortality is an outcome of interest. A complete assessment of occupational exposures would include a detailed chronological occupation history, with information on the date, duration, nature and intensity of exposures. Such an extensive assessment is not appropriate for the minimal coverage of this risk factor for control of confounding. Rather, two brief approaches have been included -- longest occupation, and checklist of known carcinogens:

A. Usual occupation will permit coding by Standard Occupation Codes, for comparison with other studies, and comparison with published data on the occupational variable most likely to have had an impact on the subject's health. It will also permit the construction of a modified Index of Social Status.

B. The checklist of exposures and occupations includes most of those for which associations with human cancer are well established (Decoufle, 1982). This will permit baseline comparability evaluations, as well as control for possible confounding by these exposures.

3.8 FAMILY HISTORY (QUESTION 30)

Susceptibility to cancer has been shown to have a genetic component for at least some cancer sites. Potential confounding by genetic factors should be considered in examining comparability between groups and in assessing outcome. The question presented here is in the format developed by the American Cancer Society for use in their nationwide Cancer Prevention Study II.

3.9 HEALTH HABITS (QUESTIONS 33-37)

Certain health practices have been shown to be associated with overall health and mortality. They may be risk factors in themselves, or markers of a physiologic state associated with impaired resistance to disease. As potential confounders, their distribution in the study groups should be examined.

A Health-Related Practices Index, based on smoking, physical activity, alcohol use, weight status, and sleep habits, was shown in a 9-year follow-up of 6,928 adults to be associated with subsequent mortality at the $p<0.001$ level (Wingard et al., 1982). This index can be constructed from questions 8, 13 (alcohol), 31, 32, 35 and 37 of the FULL87 questionnaire. The wording of the questions is either exactly that used in the study referred to above or has been modified slightly after consultation with the original authors. The DIETSYS software utilities can use these questions to calculate a "Health Practices Index."

3.10 PSYCHOLOGICAL AND SOCIAL FACTORS (QUESTIONS 38-39)

Two large-scale, long-term follow-up studies, one by Berkman and Syme (1979) in Alameda County, California (6,928 adults, 9 years) and the other by House, Robbins and Metzner in Tecumseh, Michigan (2,754 adults, 10 years) have found a strong association between psychosocial factors and subsequent mortality. "The associations of the social variables with mortality persist even with controls for age and a wide range of indicators of health and disease status. Thus, social relationships and activities appear to constitute an additional set of consequential risk factors for mortality" (House et al., 1982).

Since virtually all major intervention studies will examine total mortality for evidence of adverse intervention effects even where this is not the primary focus, the possibility of an uneven distribution of a potential confounder in the study groups should be examined.

A checklist of stressful life events is not included here, because of the potential for creating psychological distress in the respondent, and because a list covering relevant events for a wide age range of respondents would be too long to be feasible. Coping strategies or perceived locus of control versus helplessness in the face of stressful events are also not included here, since no measurement instrument was found which clearly could assess these factors in a few questions, over a wide age range of respondents, and for an extensive time period in their lives. If such measures become available, they should be given serious consideration.

Questions 38-39: The wording of question 38 is exactly that used in the Berkman Social Network Index. The wording of question 39 was modified and shortened slightly, after consultation with the original author. The DIETSYS software utilities can use these questions to calculate a "Social Network Index".