

2.0 INTRODUCTION TO THE HEALTH HABITS AND HISTORY QUESTIONNAIRE

2.1 PURPOSE

The purpose of the Health Habits and History Questionnaire (HHHQ) is to facilitate the collection of a minimum core of data which, if gathered in a standard manner, would enhance comparability between studies and facilitate other uses discussed below. The Division of Cancer Prevention and Control has an interest in enhancing the interpretability and comparability of studies, since research results, particularly in the area of diet, may find broad application in the general population. While some of the non-diet data are specific to cancer, it should be noted that the diet questionnaire is not intended to focus only on cancer risk factors, but rather to assess a broad range of dietary risk factors.

The intent is to provide a standard format for the collection of a minimum core of data on important, well-established potential confounders and predictive factors. Investigators would, of course, usually wish to design their own instrument for the primary research question; this questionnaire is not intended to replace that, but only to provide a standard format for the collection of confounders and predictors which investigators would wish to control for but not focus on. However, in view of its brief and self-administered format, investigators might consider including even those sections of the questionnaire which overlap their primary research question, for comparability with other studies. Most questions were selected, usually with the exact wording, from large-scale or continuing national studies (see Section 3, Description and Justification of HHHQ Areas of Inquiry). This provides the ability to compare results not only with other investigations but with representative national data.

2.2 SIGNIFICANCE AND JUSTIFICATION

2.2.1 Between Studies: Comparability

The lack of comparability between studies has been of concern to researchers for decades (Gordis, 1979). Standardization would facilitate the interpretation of results from different studies, by ensuring the collection of an identical set of core data on demographic and confounding variables in a uniform manner.

Each study seeks to make a contribution to knowledge. However, if several investigators study the same disease and intervention and find inconsistent results, then the real advancement of knowledge is limited. Usually all that can be said is that the study populations were different in some way, but without enough precision to be explanatory. On the other hand, if data on a range of descriptive variables are available, one can then pinpoint the differences and analyze subgroups.

For example, did the prevalence of smoking differ in the two studies? One could then look separately at the effect in smokers and nonsmokers. Perhaps the drug only works in nonsmokers. Or perhaps the drug is only effective in individuals with an inadequate diet, or only in those whose lifestyle promotes resistance to physiologic insult. If data on such factors are available, conflicting data may be resolvable and the value of any single study is thereby enhanced.

2.2.2 Within Studies: Critical Confounders

Occasionally, data on important potentially confounding variables are not collected, or are not reported in all investigations. Use of the HHHQ questionnaire will ensure that at least a minimum set of important confounding or explanatory variables is collected and available for analysis, thus improving the validity and interpretability of results.

Our knowledge of the full sequence of biochemical and physiologic interactions involved in an association or an intervention is incomplete. Given these gaps in our knowledge about interactions, it is important to collect data about risk factors with a broader range of impact than simply well-demonstrated risk factors for the disease in question. For example, smoking might not be a proven risk factor for the disease under study. However, smoking is known to have wide-ranging physiologic effects, and could potentially modulate the individual's response to an intervention or to other exposures. Similarly, a number of areas such as diet, lifestyle or socioeconomic factors have a clear relationship to disease incidence, though the mechanism may be as yet unclear. Unless the data on known potent modulators of health are collected, the results of any given study will not be as strong as they could have been.

In investigations of potential protective factors, and especially in intervention trials, there is an additional reason to examine confounding factors beyond those limited to the disease in question. That is, in intervention trials we must balance risks and benefits; we must examine the effect of the intervention not just on the disease under study but on a wide range of other health outcomes. In the past, some investigators have been faced with an intervention which is apparently effective against the target disease, but either does not reduce or actually increases total all-cause mortality. Even within a single broad classification of disease, such as cancer, an intervention might reduce incidence at the target cancer site, while increasing incidence at another site.

Thus, it is important for trials to include data which will permit careful interpretation of apparently adverse effects. Suppose, for example, that one sees an apparent excess of deaths from cardiovascular disease in the intervention group. One would be very reluctant to recommend the intervention for large-scale use in the general population, and an effective drug might have to be abandoned. If appropriate risk factor data were available, however, one could control for potential confounding due to those factors. This would permit the investigators either to confirm that the risk persists after controlling for confounding, or to reassure policymakers and the public that the apparent adverse effect was due not to the intervention but to other risk factors such as smoking.

Thus, the interpretability of each study will be enhanced by the collection of data not just on risk factors for the target disease, but also risk factors relevant to a broader spectrum of outcomes.

2.2.3 Prospective Uses

Many investigators may wish to monitor long-term health and mortality outcomes. This is particularly true for studies carried out under the auspices of the Division of Cancer Prevention and Control, which all have as their underlying purpose the development of methods for the prevention and control of cancer in the general population. Methods which are found to be effective in intervention trials may ultimately be introduced to the general population on a very large scale and over time periods which could be as long as an individual's lifetime. Since this

is the potential result of these investigations, ethical as well as scientific considerations make it essential to monitor the effects of these interventions over a very long time scale. Furthermore, as discussed under 1.2.1 and 1.2.2 above, it is necessary to evaluate any long-term effects in the light of knowledge about potentially confounding factors. The collection of the necessary data for long-term monitoring is thus important to the mission of this Division. The Health Habits and History Questionnaire permits the collection of the necessary information in a standardized way.

Furthermore, because the use of the Health Habits and History Questionnaire entails the collection of a standardized set of information, it is hoped that it will permit the investigation of risk factors and protective factors, associations with outcome, and prevention strategies in situations now difficult to investigate because of the long time periods and large numbers involved. For example, the health effects of different dietary intakes may be difficult to study prospectively over a period of three to five years because of low incidence rates and consequent low statistical power. Long-term vital status monitoring or appropriate pooling of data may permit the investigation of such factors.

Thus, it is hoped that the use of this standardized data collection instrument will enhance the interpretation of individual studies, improve comparability between studies, and permit long-term prospective use and pooling of data.

2.3 AVAILABLE DATA MANAGEMENT AND ANALYTIC TOOLS

Three versions of the questionnaire have been included in this packet. The full questionnaire (including both non-diet and diet sections) takes 35-40 minutes to self administer. Coding takes approximately 15 minutes and does not require any special expertise. The data management and analytic tools associated with the HHHQ include:

- A. Instructions for self-administration (Section 5 and Section 5, Appendix A).
- B. Instructions for interviewers (Section 5 and Section 5, Appendix B).
- C. A diskette containing the Dietary Analysis Personal Computer System (DIETSYS) has been included. The user documentation for DIETSYS is included in this document.
- D. Suggestions for data management and analysis (Section 5).

The following are available upon request:

- E. Very brief screening questionnaires (for fat, fruit/vegetable/fiber and for calcium intake).
- F. HHHQ translations into Spanish and Italian.
- G. Training video for interviewers (available to borrow).

- H. The codebook for Version 01, February, 1985 questionnaires. This is available for investigators who have already collected data with this earliest version of the HHHQ. Other investigators should use one of the three current NCI questionnaires.
- I. Portion size database for children 1-2 years and 3-4 years of age.

To receive item E, F or I, please send a written request to the following address:

Dr. Gladys Block
419 Warren Hall
University of California
Berkeley, CA 94720

To receive item G please send a written request to the following address:

Ms. Anne Hartman
National Cancer Institute
EPN 313
9000 Rockville Pike
Bethesda, MD 20892

To receive item H please send a written request to either address.