
National Survey of Family Growth

Cycle 5: 1995

Omitted Items File Documentation

Supplement to the 1995 National Survey of Family Growth Public Use Data
Files and Documentation

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
June, 1997

Omitted Items File Documentation

Supplement to the 1995 National Survey of Family Growth

Introduction

Interviewing for the National Survey of Family Growth (NSFG), Cycle 5, was conducted in January through October of 1995 by Research Triangle Institute, Inc., under contract to the National Center for Health Statistics (NCHS). In-person interviews were conducted with 10,847 women 15-44 years of age, of all marital statuses. The entire survey was administered using computer-assisted personal interviewing (CAPI). A portion of the survey, including some of the more sensitive items, was self-administered by the respondent using audio computer-assisted self-interviewing (Audio CASI or A-CASI). In Audio CASI, the respondent hears the questions through headphones and enters her own answers into the laptop computer.

In March 1997, NCHS released public use data files containing most of the information collected in the 1995 NSFG. Because of its legal responsibilities to protect the confidentiality of the personal data that respondents have voluntarily entrusted to it, NCHS has deemed it necessary and prudent to omit certain items from the public use data files. The omitted items file includes *all* of the Audio CASI questions, as well as selected items from public use respondent-based data file.

In Audio CASI respondents were asked questions pertaining to each of up to 10 abortions (date, length of pregnancy, contraception use before terminated pregnancy), number of sexual partners, forced sexual intercourse, HIV risk behaviors, and childhood family violence and drinking/drug use. Respondents were also asked which interview mode (e.g., interviewer-administered versus self-administered) they preferred for sensitive questions. Many of these Audio CASI items are analogous to variables included in the interviewer-administered portion of the survey (available on the public use file), because one of the main objectives of Audio CASI was to give respondents an additional chance to report sensitive information. In addition to the data collected through Audio CASI, the omitted items file includes information on the respondents' experience with sexually transmitted diseases.

While these items were not made generally available on the public use files, NCHS wishes to promote the scientific purposes for which the data were collected. To this end, NCHS has made the data available to researchers for specified scientific analyses, under special arrangements that assure confidentiality and protection of the data commensurate with that provided by the Center itself. This Omitted Items File Documentation Supplement describes the data file containing the items that have been omitted from the 1995 NSFG public use files.

The omitted items data file is intended to be used in conjunction with the NSFG public use data files. The user is referred to the User's Guide for the 1995 NSFG public use files for a detailed description of the data preparation procedures, sample design, and coding conventions that apply not only to the public use files, but to the omitted items file as well.

In addition to the Audio CASI data and other items described above, the omitted items file includes date of interview and time stamps for each section of the interview. With rare exception, all interviews were completed on the same day; A_DOI (available on the public use files) will almost always be the same as other section start dates. Time stamps (i.e., time when respondent began a particular section) were included at key points in the CAPI program to measure interview length.

File Characteristics and Data Layout

There are 2 ASCII data files containing the omitted items for the 1995 National Survey of Family Growth. One file is respondent-based and contains 1 record for each of the 10,847 women interviewed. For the convenience of users, we have provided a second file that is abortion-based (or interval-based), with 1 record for each abortion reported by respondents in Audio CASI.

Respondent-Based File:

compressed data set name:	NSFGOM95.ZIP
uncompressed data set name:	NSFGOM95.DAT
number of records:	10,847 (respondents interviewed in 1995 NSFG)
record length:	1,035
compressed size:	1,921,763 bytes (spans 2 diskettes)
uncompressed size:	11,248,339 bytes

Note: PKUNZIP.EXE is provided to uncompress the data set.

Abortion-Based File:

data set name:	ABORTINT.DAT
number of records:	3,421 (abortions reported in A-CASI)
record length:	54
size:	191,576 bytes

Both of these data files include CASEID, the case identification number that enables users to merge omitted file data with data from the public use files.

File Layout for Respondent-Based File:

<u>Column Locations</u>	<u>Variable(s)</u>
1-8	CASEID
9-10	Blank
11-494	A-CASI variables (Section J of main interview)
495-499	STD variables (omitted from Section H of main interview)
500-1035	Dates of interview & Time stamps

File Layout for Abortion-Based File:

<u>Column Locations</u>	<u>Variable(s)</u>
1-5	UNIQ_ID (unique abortion file ID number)
6-13	CASEID
14-15	ABORT_NO (abortion number)
16-54	Abortion-specific data from A-CASI

Key to Codebook Documentation

For each item, the codebook documentation includes a variable name, numbers indicating the beginning and ending column locations on the data file, the question text for the item, the “inapplicable specification,” the numerical values into which responses were coded with detailed value labels, and unweighted frequencies.

Variable name:

The variable name corresponds exactly or approximately to the question name that appears in the CAPI Reference Questionnaire, which is the readable version of the interview program. (See “Key to Respondent and Interval File indexes” in Public Use File User’s Guide for further details on variable names.)

Question text:

The question text is given here, and where appropriate, CAPI wording variants are presented in parentheses.

Inapplicable specifications:

Some questions were asked of all respondents, and these are indicated as such. But for many variables in the codebook, beneath the question text is a statement specifying the universe of respondents **who was not asked** the question. The “inapplicable specification” gives a precise description of all the conditions under which respondents were **not** asked the question -- that is, cases for whom the question was inapplicable.

Response categories and unweighted frequencies:

For categorical variables and several continuous variables, the documentation lists all possible values with descriptive value labels. Most century month (date) variables and continuous variables have been collapsed into meaningful, yet manageable, groups. Frequencies are shown for the total sample of women. (In the public use file codebooks, the numbers of respondents in each response category are also shown separately for Hispanic women, non-Hispanic white women, and non-Hispanic black women.) These marginal distributions are raw unweighted counts of sample women or abortion intervals; they are not estimates of numbers or distributions for the population represented by the sample. **To make national estimates, sample cases must be weighted** (see section on “Organization and use of the data files” in the Public Use File User’s Guide). The unweighted marginal distributions are provided only as a guide to the researcher, who often wants to know in advance if the number of cases in a particular category is large enough to sustain a particular analysis.