1996 National Health Interview Survey (NHIS)
Access to Care Final Public Use Data File

1. The 1996 National Health Interview Survey (NHIS) Access data file contains a variety of data items addressing access to health care services. These data items are identical to the 1995 Access data items.

   As a result of a Federal government furlough, two weeks of data collection were omitted in January of 1996. In addition, in order to test the changing NHIS core questionnaire, for much of the year the sample was split between the old (paper) and the new (computerized) versions of the core questionnaire. This data file includes only data obtained from the paper version of the NHIS questionnaire. The sample size is considerably smaller than in the previous year (63,402 vs. 102,467). The weights have been adjusted for these factors to produce national estimates, however, there may be a minor seasonal effect which is not corrected by the weighting.

2. The 1996 Access supplement was administered for the full year (except as mentioned above) in all of the NHIS sample households interviewed with the old paper core using three-quarters of the sample households from January-June and one-half of the sample households from July-December. Information was collected from a household respondent about all family members who participated in the NHIS.

   The 1996 Access file is structured in the following way:

   a. The NHIS person record from the core questionnaire (locations 1-335)
   b. The weight fields (locations 201-236)
   c. Fields needed for calculating variances (locations 337-358)
   d. Data from supplement (locations 401-439)

   Note: All data from the Access supplement have been shifted to start in location 401 in order to accommodate a longer public use person record required by the new sample design in 1995.

3. In 1996, two types of item non-response were identified:

   (1) "Not ascertained" (codes 8, 98, or 998) includes blanks when there should have been a response or when an impossible code appeared and;
   (2) Responses of "don't know" or "refused" when the question was asked (codes 9, 99 and 999).

4. The overall response rate for the 1996 Access was 90.0 percent. This response rate was calculated as follows:

> Household response rate from core of 93.8 percent multiplied by 95.9 percent who responded to the Access section yields an overall response rate of 90.0 percent.

Dummy records were created for those with no response to the entire section (see file location 400).

5. Weights and variances:

Since the NHIS uses a multistage sample design to represent the civilian non-institutionalized population of the United States, weights must be used to make accurate estimates based on data from the National Health Interview Survey.

A set of weights are included on the 1996 file:

The first weight listed below (i.e. the Final Basic Weight) will be used in most analyses of the Access data.

The Final Basic Weight (location 219-227) is the equivalent of the Annual Final Basic Weight found on the NHIS Person Record of the Basic Health and Demographic component of the survey (i.e. the Core questionnaire). A national estimate of all person level variables can be made using this weight.

This weight will be used in conjunction with Access data items in file locations 401-439.

The Final Quarter Basic Weight before age-sex-race/ethnicity adjustment (loc. 172-177) is required by some software packages for variance estimation for surveys with complex sample designs. This weight is also included on the file.

As mentioned above, the sample design for the NHIS was changed for 1995. Data from 1995 and 1996 can be combined with data from previous years, however, variances for 1995 and 1996 must be calculated separately from variances of previous years. In addition, because of the smaller sample size in 1996, some of the design elements were combined for reasons of confidentiality. The exact changes are included in the explanation below.

There are a number of computer programs that yield variance estimates for data based on complex sample surveys. Some are based on replication approaches and others are based on Taylor

linearization approaches. In addition to the Final Quarter Basic Weight before age-sex-race/ethnicity adjustment (which is the weight prior to post-stratification), included on the Access file is the substratum for variance estimation (loc. 342-343), the secondary sampling unit (loc. 344-350), Panel 4 (loc. 352), the variance PSU for 1996 (loc. 358), the collapsed variance stratum for 1996 (loc. 354-357), and the NSR Status variable for 1996 (loc. 353) to permit the analyst the capability of using such variance estimation procedures. These variables and weights are necessary for directly calculating sampling variances.

6. Estimating annual numbers of events or conditions

a. To reduce respondent error, the recall period for questions about some events is limited to two weeks. These events are: bed days and other restricted activity days, work loss and school loss days, and doctor visits. The two-week variables are found in locations 98-107 and 120-121. Estimates of the total number of occurrences of these events in the population can be derived as follows:

Number of events x 26 (number of two-week periods in a year) x Final Basic Weight

= Total number of events occurring in the population during the annual period, i.e. 1996.

Example: Number of bed days (Loc. 100-101) x 26 x Final Basic Weight (Loc. 219-227) = total number of bed days reported for the population in 1996.

b. The recall period for acute incidence conditions is also two weeks and a national estimate of the total number of acute incidence conditions is calculated using the same procedures as for two-week events for the annual period.

Number of acute incidence conditions x 26 x Final Basic Weight

= Total number of acute incidence conditions occurring in the population during 1996.

Note: An acute incidence condition is an acute condition with onset during the two weeks preceding the date of interview.

c. The recall period for information on hospitalizations is 12 months. However, in calculating number of discharges (Locations 132-133, 137-138), only discharges occurring in the past 6 months are counted. Therefore, the weighted estimates must be calculated as follows:

Number of discharges x 2 x Final Basic Weight

= Total number of discharges occurring in the population in 1996.

7. Calculation of rates for events and conditions:

The number of events or conditions estimated for the population, as described in item 6, above, can be used as the basis for calculating rates of occurrence of these events (or conditions) per person and per 100 persons for the total U.S. population and for various population subgroups.

Note: Only rates can be estimated from these data. The percent of the population experiencing a particular type of event during the data year cannot be estimated. (The percent of the population experiencing the event in the reporting period (i.e. two weeks or 6 months) can be estimated but is generally not meaningful.)

8. Data on hospital episodes and days, based on a 12-month recall are in locations 122-131. The Final Basic Weight is used for calculating estimates of these events in the same way it is used for all other person-based variables. These variables do permit estimating the percent of the population in this annual period experiencing a hospital episode in the past year and the percent of that population having a specified number of hospital days.

9. Guidelines for Citation of Data

With the goal of mutual benefit, the National Center for Health Statistics (NCHS) requests that recipients of data files cooperate in certain actions related to their use.

Any published material derived from the data should acknowledge NCHS as the original source. The suggested citation to appear at the bottom of all tables is as follows:

> Source: National Center for Health Statistics (1996)

When cited in a bibliography, the suggested citation should read:

> National Center for Health Statistics (1998). Data File Documentation, National Health Interview Survey of Access to Care, 1996 (machine readable data file and documentation), National Center for Health Statistics, Hyattsville, Maryland.

The published material should also include a disclaimer that credits any analyses, interpretations, or conclusions reached to the author (recipient of the data file) and not to NCHS, which is responsible only for the initial data. Consumers who wish to publish a technical description of the data should make a reasonable effort to insure that the description is not inconsistent with that published by NCHS.

1996 NATIONAL HEALTH INTERVIEW SURVEY (NHIS)
ACCESS TO CARE PUBLIC USE FILE

Outline of Items and Codes

63,402 Records

File Locations	Item No.	Frequency	Items and Codes
1-2	-		RECORD TYPE
		63,402	70. Access to Care Supplement
3-4	HH-2		PROCESSING YEAR
		63,402	96. 1996
5-14	Recode	-	HOUSEHOLD ID
15-16	-	-	PERSON NUMBER
17-18	-	-	BLANK (Record Serial Number on some other record types)
19-20	-		SAMPLING WEEK CODE (Numbered within quarter)
		3,680	01. Week 1
		3,422	02. Week 2
		5,414	03. Week 3
		4,966	04. Week 4
		5,009	05. Week 5
		4,867	06. Week 6
		5,512	07. Week 7
		5,183	08. Week 8
		5,390	09. Week 9
		4,992	10. Week 10
		4,714	11. Week 11
		5,128	12. Week 12
		5,125	13. Week 13

1996 NHIS ACCESS TO CARE PUBLIC USE FILE

File Locations	Item No.	Frequency	Items and Codes
21	Recode		LATE INTERVIEW (OR LAST ATTEMPT) FLAG
		36,891	0. Interview not late
		18,591	1. One week late
		6,372	2. Two weeks late
		1,548	3. Unknown
22-23	HH-11c,d		TYPE OF LIVING QUARTERS:
			Housing Unit = (00-07)
		1,234	00. Housing unit; kind unknown
		58,363	01. House, apartment, flat
		37	02. HU in nontransient hotel, motel, etc.
		5	03. HU-permanent in transient hotel, motel, etc.
		20	04. HU in rooming house
		3,072	05. Mobile home or trailer with no permanent room added
		450	06. Mobile home or trailer with one or more permanent rooms added
		51	07. HU not specified above
			Other Unit = (08-13)
		39	08. Quarters not HU in rooming or boarding house
		3	09. Unit not permanent in transient hotel, motel, etc.
		14	10. Unoccupied site for mobile home, trailer, or tent
		90	11. Student quarters in college dormitory
		24	12. Other unit not specified above
		0	13. Other unit; kind unknown
24	HH-12a		HAS TELEPHONE
		57,080	1. Yes, phone number given
		2,558	2. Yes, no phone number given
		3,099	3. No
		665	4. Unknown
25	A-1		SEX
		30,358	1. Male
		33,044	2. Female

1996 NHIS ACCESS TO CARE PUBLIC USE FILE

File Locations	Item No.	Frequency	Items and Codes
26	-		AGE IMPUTED FLAG
		63,400	0. Age known
		2	1. Age unknown, imputed as 34
27-28	Person Column		AGE
		976	00. Under 1 year
		62,241	01-89. Number of years
		185	90. 90+ years old
29	Recode		AGE RECODE #1
		4,918	1. Under 5 years
		13,210	2. 5-17 years
		5,568	3. 18-24 years
		19,974	4. 25-44 years
		12,598	5. 45-64 years
		2,213	6. 65-69 years
		1,928	7. 70-74 years
		2,993	8. 75 years and over
30	Recode		AGE RECODE #2
		6,008	1. Under 6 years
		11,219	2. 6-16 years
		6,469	3. 17-24 years
		9,603	4. 25-34 years
		10,371	5. 35-44 years
		7,673	6. 45-54 years
		4,925	7. 55-64 years
		4,141	8. 65-74 years
		2,993	9. 75 years and over
31-32	Recode		AGE RECODE #3
		2,905	00-35. Months
		60,497	36. Over 3 years old
33	-		MONTH OF BIRTH IMPUTED FLAG
		60,272	0. Month known
		3,090	1. Month unknown, '8' imputed
		40	9. Date of birth unknown

File Locations	Item No.	Frequency	Items and Codes
34-39	A-3	-	MONTH AND YEAR OF BIRTH
34-35			Month 01. January 08. August 02. February 09. September 03. March 10. October 04. April 11. November 05. May 12. December 06. June 99. Unknown 07. July
36-39			Year of Birth
		168	1905. 1905 and before
		63,194	1906-1997. 1906-1997
		40	9999. Unknown
40	Recode		HISPANIC ORIGIN IMPUTED FLAG
		62,712	0. Hispanic origin known
		690	1. Hispanic origin imputed from reference person
41-42	A-6		MAIN RACIAL BACKGROUND* - Reported (see notation for locations 43-45)
		46,996	01. White
		9,027	02. Black/African American**
		482	03. Indian (American)
		325	06. Chinese
		455	07. Filipino
		987	15. Other API (includes Hawaiian, Korean, Vietnamese, Japanese, Asian Indian, Samoan, and Guamanian)
		3,797	16. Other race (includes Eskimo and Aleut)
		181	17. Multiple race
		1,152	99. Unknown

* Some categories may be too small to analyze separately and therefore may produce unreliable estimates; in addition, counts may not agree with those produced by the Census Bureau.

** For convenience, the category 'Black/African American' will be shown as 'Black' in all observed race or race recode locations throughout the documentation.

1996 NHIS ACCESS TO CARE PUBLIC USE FILE

File Locations	Item No.	Frequency	Items and Codes	
43-45	Recode		RACE RECODES	
43			Recode 1*	Persons whose Main
		51,962	1. White	Racial Background
		9,065	2. Black	(location 41-42) was
		2,375	3. Other	"other" or "unknown"
				were classified in
44			Recode 2	the following recodes
				by using the racial
		51,962	1. White	background observed
		11,440	2. Non-white	by the interviewer.
				Use of these recodes
45			Recode 3	is recommended for
				estimating statistics
		9,065	1. Black	for the groups shown
		54,337	2. Non-black	here.
46-47	A-5		HISPANIC ORIGIN**	
		12	00. Multiple Hispanic	
		1,387	01. Puerto Rican	
		646	02. Cuban	
		3,935	03. Mexican-Mexicano	
		4,391	04. Mexican-American	
		99	05. Chicano	
		1,133	06. Other Latin American	
		1,184	07. Other Spanish	
		338	08. Spanish, DK type	
		114	09. Unknown if Spanish origin	
		50,163	10. Not Spanish origin	
48	L-7		MARITAL STATUS	
		14,293	0. Under 14 years old	
		28,314	1. Married - spouse in household	
		494	2. Married - spouse not in household	
		3,022	3. Widowed	
		3,475	4. Divorced	
		1,102	5. Separated	
		12,070	6. Never married	
		632	7. Unknown	

*This recode is used to define race in the Current Estimates tables.

**If unknown, the family reference person code was imputed. A flag indicating imputation is in loc. 40 and the relationship to reference person is in loc. 63.

File Locations	Item No.	Frequency	Items and Codes
49	L-1		VETERAN STATUS
		38,331	1. Non-veteran
		1,311	2. WW I and WW II
		800	3. Korean War
		1,715	4. Vietnam veteran
		810	5. Post-Vietnam
		1,070	6. Other service
		163	7. Served in Armed Forces, unknown if war veteran
		1,074	8. Unknown if served in Armed Forces
		18,128	Blank. Under 18 years old
50	L-1		ACTIVE GUARD/RESERVE STATUS FOR PERSONS ON ACTIVE DUTY IN ARMED FORCES
		38,331	0. Non-veteran
		363	1. All service in Guard/Reserve
		740	2. Some service in Guard/Reserve
		36	3. Unknown if all service in Guard/Reserve
		4,159	4. No active service in Guard/ Reserve
		1,645	5. Unknown if ever active member in Guard/Reserve or served in Armed Forces
		18,128	Blank. Under 18 years old
51-52	L-2		EDUCATION OF INDIVIDUAL - COMPLETED YEARS
		2,628	00. Never attended; kindergarten only
		19,974	01-11. Grades 1-11
		16,207	12. Grade 12
			College:
		3,321	13. 1 year
		4,433	14. 2 years
		1,602	15. 3 years
		5,424	16. 4 years
		1,039	17. 5 years
		2,762	18. 6 years or more
		1,094	19. Unknown
		4,918	Blank. Under 5 years old

File Locations	Item No.	Frequency	Items and Codes
53	Recode		EDUCATION OF INDIVIDUAL
		2,628	0. None; kindergarten only
		12,480	1. 1-8 years (elementary)
		7,494	2. 9-11 years (high school)
		16,207	3. 12 years (high school graduate)
		9,356	4. 1-3 years (college)
		5,424	5. 4 years (college graduate)
		3,801	6. 5+ years (post-college)
		1,094	7. Unknown
		4,918	Blank. Under 5 years old
54-55	-		HIGHEST EDUCATION OF RESPONSIBLE ADULT FAMILY MEMBER (Detail)
		151	00. Never attended; kindergarten only
		9,278	01-11. Grades 1-11
		20,685	12. Grade 12
			College:
		4,719	13. 1 year
		7,364	14. 2 years
		2,688	15. 3 years
		9,589	16. 4 years
		2,073	17. 5 years
		6,048	18. 6 years or more
		807	19. Unknown
56	Recode		HIGHEST EDUCATION OF RESPONSIBLE ADULT FAMILY MEMBER
		151	0. None; kindergarten only
		3,807	1. 1-8 years (elementary)
		5,471	2. 9-11 years (high school)
		20,685	3. 12 years (high school graduate)
		14,771	4. 1-3 years (college)
		9,589	5. 4 years (college graduate)
		8,121	6. 5+ years (post-college)
		807	7. Unknown

File Locations	Item No.	Frequency	Items and Codes
57	L-8		FAMILY INCOME $20,000 OR LESS
		18,637	1. Less than $20,000
		42,177	2. $20,000 or more
		2,588	3. Unknown
58-59	L-8		FAMILY INCOME
		243	00. Less than $1,000
		366	01. $ 1,000 - $1,999
		319	02. 2,000 - 2,999
		356	03. 3,000 - 3,999
		433	04. 4,000 - 4,999
		694	05. 5,000 - 5,999
		695	06. 6,000 - 6,999
		699	07. 7,000 - 7,999
		701	08. 8,000 - 8,999
		941	09. 9,000 - 9,999
		1,132	10. 10,000 - 10,999
		705	11. 11,000 - 11,999
		1,304	12. 12,000 - 12,999
		838	13. 13,000 - 13,999
		828	14. 14,000 - 14,999
		1,100	15. 15,000 - 15,999
		777	16. 16,000 - 16,999
		810	17. 17,000 - 17,999
		1,042	18. 18,000 - 18,999
		1,216	19. 19,000 - 19,999
		4,786	20. 20,000 - 24,999
		4,130	21. 25,000 - 29,999
		4,150	22. 30,000 - 34,999
		3,179	23. 35,000 - 39,999
		3,180	24. 40,000 - 44,999
		2,824	25. 45,000 - 49,999
		15,137	26. $50,000 and over
		10,817	27. Unknown

1996 NHIS ACCESS TO CARE PUBLIC USE FILE

File Locations	Item No.	Frequency	Items and Codes
60	Recode		FAMILY INCOME
		1,717	0. Under $5,000
		1,389	1. $ 5,000 - $ 6,999
		2,341	2. 7,000 - 9,999
		4,807	3. 10,000 - 14,999
		4,945	4. 15,000 - 19,999
		4,786	5. 20,000 - 24,999
		8,280	6. 25,000 - 34,999
		9,183	7. 35,000 - 49,999
		15,137	8. $50,000 or more
		10,817	9. Unknown
61	Generated		NHIS POVERTY INDEX*
		48,720	1. At or above poverty threshold
		8,572	2. Below poverty threshold
		6,110	3. Unknown
62-63			FAMILY RELATIONSHIP
62	A-2		Type of Family
		6,440	&. Primary individual
		684	-. Secondary individual
		56,152	0. Primary family
		126	1-9. Secondary family
63	A-2		Relationship to Reference Person
		6,048	&. Reference person, living alone
		18,770	0. Reference person, 2+ persons in household
		13,756	1. Spouse, other spouse NOT in Armed Forces and living at home
		151	2. Spouse, other spouse IN Armed Forces and living at home
		21,000	3. Child of reference person or spouse
		1,252	4. Grandchild of reference person or spouse
		645	5. Parent of reference person or spouse
		1,763	6. Other relative
		17	7. Child of military family with no eligible reference person
		0	9. DK or refused

*Based on family size, number of children under 18 years old and family income using the 1995 poverty levels derived from the August, 1996 Current Population Survey.

File Locations	Item No.	Frequency	Items and Codes
64	Recode		FAMILY RELATIONSHIP
		6,048	1. Living alone
		1,076	2. Living only with non-relative
		28,297	3. Living with spouse
		27,981	4. Living with relative - other
65-66	Generated	-	SIZE OF FAMILY*
			Unrelated individuals are coded 01
67	Generated		SIZE OF FAMILY RECODE
		62,665	1-8. Number of members
		737	9. 9+ members
68	A-2		PARENT/OTHER ADULT RELATIVE (under 25 years old and never married)
		13,233	1. Both parents, no other relative
		3,521	2. Mother only
		324	3. Father only
		1,835	4. Both parents and other 21+ year old adult relative
		1,271	5. Mother and other 21+ year old adult relative
		152	6. Father and other 21+ year old adult relative
		281	7. No parent, but one 21+ year old adult relative
		427	8. No parent, but two or more 21+ year old adult relatives
		164	9. Unknown
		849	0. Other
		41,345	Blank. Not applicable (25+ years old or ever married)

*Count includes spouse in military but living at home.

File Locations	Item No.	Frequency	Items and Codes
69	B-1 B-8		MAJOR ACTIVITY (18+ years old)
		28,283	1. Working
		7,516	2. Keeping house
		2,504	3. Going to school
		6,513	4. Something else
		458	5. Unknown
		18,128	Blank. Not applicable (under 18 years old)
70	G-4		HEALTH STATUS
		22,720	1. Excellent
		18,116	2. Very Good
		15,237	3. Good
		4,899	4. Fair
		1,697	5. Poor
		733	6. Unknown
71	Recode		ACTIVITY LIMITATION STATUS*- (all ages)
		3,018	1. Unable to perform major activity
		3,343	2. Limited in kind/amount major activity
		2,727	3. Limited in other activities
		54,314	4. Not limited (includes unknowns)
72	Recode		ACTIVITY LIMITATION STATUS MEASURED BY "ABILITY TO WORK" (18-69 years old)
		2,829	1. Unable to work
		1,765	2. Limited in kind/amount of work
		1,490	3. Limited in other activities
		34,269	4. Not limited (includes unknowns)
		23,049	Blank. Not applicable (under 18 years old, 70+ years old)

*This location is used to categorize persons with limitation of activity in the Current Estimates tables.

File Locations	Item No.	Frequency	Items and Codes
73	B-11		LIMITATION OF SCHOOL ACTIVITIES (5-17 years old)
		81	1. Unable to attend school
		438	2. Attends special school/classes
		74	3. Needs special school/classes
		124	4. Limited in school attendance
		282	5. Limited in other activities
		12,211	6. Not limited (includes unknowns)
		50,192	Blank. Not applicable (under 5 years old or 18+ years old)
74	B-14		NEEDS HELP WITH PERSONAL CARE (5-59 years old and limited, or 60-69 years old)
		408	1. Unable to perform personal care needs
		822	2. Limited in performing other routine needs
		8,488	3. Not limited in performing personal or routine needs
		404	4. Unknown
		53,280	Blank. Not applicable (under 5 years old; 5-59 years old not limited; 70+ years old)

File Locations	Item No.	Frequency	Items and Codes
75	D-1		EMPLOYMENT STATUS IN PAST 2 WEEKS (18+ years old)
			In the Labor Force: (1-7)
			Currently employed: (1-3)
		28,767	1. Worked in past 2 weeks
		494	2. Did not work, has job; not on lay-off and not looking for work
		28	3. Did not work, has job; looking for work
			Unemployed: (4-7)
		93	4. Did not work, has job; on lay-off
		6	5. Did not work, has job; on lay-off and looking for work
		200	6. Did not work, has job; unknown if looking or on lay-off
		960	7. Did not work, has no job; looking for work or on lay-off
		14,726	8. Not in Labor Force (18+ years old)
		18,128	Blank. Not applicable (under 18 years old)

File Locations	Item No.	Frequency	Items and Codes
76	L-6		CLASS OF WORKER
		14,726	0. Not in labor force
		21,596	1. Private company
		756	2. Federal Government employee
		1,388	3. State Government employee
		2,369	4. Local Government employee
		789	5. Incorporated business
		2,615	6. Self-employed
		29	7. Without pay
		35	8. Never worked
		971	9. Unknown
		18,128	Blank. Under 18 years old
77-79	-		BLANK
80-81	Recode	-	INDUSTRY RECODE 1 SEE APPENDIX B
82-83	Recode	-	INDUSTRY RECODE 2 SEE APPENDIX B
84-86	-		BLANK
87-88	Recode	-	OCCUPATION RECODE 1 SEE APPENDIX C
89-90	Recode	-	OCCUPATION RECODE 2 SEE APPENDIX C

File Locations	Item No.	Frequency	Items and Codes
91	L-R		RESPONDENT FOR CORE QUESTIONS
		27,761	1. Self (entirely)
		3,242	2. Self (partly)
		31,712	3. Proxy
		687	4. Unknown
92	Recode		CONDITION LIST ASSIGNED AND ASKED
		10,548	1. Condition List 1, Skin and musculoskeletal
		10,568	2. Condition List 2, Impairments
		10,388	3. Condition List 3, Digestive
		10,472	4. Condition List 4, Miscellaneous
		10,113	5. Condition List 5, Circulatory
		10,575	6. Condition List 6, Respiratory
		738	7. Unknown
93-94	G-5		HEIGHT WITHOUT SHOES (18+ years old)
		363	58. 58 inches or less
		43,782	59-76. Number of inches
		246	77. 77 inches or more
		883	99. Unknown
		18,128	Blank. Under 18 years old
95-97	G-5		WEIGHT WITHOUT SHOES (18+ years old)
		288	097. 97 pounds or less
		42,919	098-289. Number of pounds
		320	290. 290 pounds or more
		1,747	999. Unknown
		18,128	Blank. Under 18 years old
98-99	Recode		TOTAL RESTRICTED ACTIVITY DAYS IN PAST TWO WEEKS
		57,092	00. None
		6,310	01-14. Days
100-101	D-4		BED DAYS IN PAST TWO WEEKS
		59,733	00. None
		3,669	01-14. Days

File Locations	Item No.	Frequency	Items and Codes
102-103	D-2		WORK-LOSS DAYS IN PAST TWO WEEKS
		61,810	00. None
		1,592	01-14. Days
104-105	D-3		SCHOOL-LOSS DAYS IN PAST TWO WEEKS
		62,431	00. None
		971	01-14. Days
106-107	D-6		OTHER DAYS OF RESTRICTED ACTIVITY IN PAST TWO WEEKS
		60,449	00. None
		2,953	01-14. Days
108-110	G-2		BED DAYS IN PAST 12 MONTHS
		35,920	000. None
		26,460	001-365. 1-365 days
		1,022	366. Unknown
111	Recode		BED DAYS IN PAST 12 MONTHS
		35,920	0. None
		20,817	1. 1-7 days
		4,084	2. 8-30 days
		1,243	3. 31-180 days
		316	4. 181-365 days
		1,022	5. Unknown
112-114	G-3		DOCTOR VISITS IN PAST 12 MONTHS
		15,335	000. None
		47,627	001-996. Visits
		0	997. 997+ visits
		440	998. Unknown
115	G-3		INTERVAL SINCE LAST DOCTOR VISIT
		182	0. Never
		48,478	1. Less than 1 year
		6,010	2. 1 to less than 2 years
		4,842	3. 2 to less than 5 years
		2,128	4. 5 years or more
		1,762	5. Unknown

File Locations	Item No.	Frequency	Items and Codes
116-117	Generated	-	NUMBER OF CONDITIONS
118-119	Generated	-	NUMBER OF ACUTE INCIDENCE CONDITIONS
120-121	Generated	-	NUMBER OF TWO-WEEK DOCTOR VISITS
122-123	Generated	-	NUMBER OF SHORT-STAY HOSPITAL EPISODES IN PAST 12 MONTHS
124-126	Generated	-	SHORT-STAY HOSPITAL EPISODE DAYS IN PAST 12 MONTHS
127-128	Generated	-	NUMBER OF SHORT-STAY HOSPITAL EPISODES IN PAST 12 MONTHS EXCLUDING DELIVERY*
129-131	Generated	-	SHORT-STAY HOSPITAL EPISODE DAYS IN PAST 12 MONTHS EXCLUDING DELIVERY*
132-133	Generated	-	NUMBER OF SHORT-STAY HOSPITAL DISCHARGES IN PAST 6 MONTHS
134-136	Generated	-	NUMBER OF DAYS IN SHORT-STAY HOSPITAL IN PAST 12 MONTHS FOR DISCHARGES IN PAST 6 MONTHS
137-138	Generated	-	NUMBER OF SHORT-STAY HOSPITAL DISCHARGES IN PAST 6 MONTHS EXCLUDING DELIVERY*
139-141	Generated	-	NUMBER OF DAYS IN SHORT-STAY HOSPITAL IN PAST 12 MONTHS FOR DISCHARGES IN PAST 6 MONTHS EXCLUDING DELIVERY*

*Based on operation codes and reason entered hospital

File Locations	Item No.	Frequency	Items and Codes
142-143	-		BLANK
144	L-9b		YEARS LIVED IN STATE OF PRESENT RESIDENCE
		1,766	1. Less than 1 year
		5,762	2. 1 yr., less than 5 yrs.
		6,429	3. 5 yrs., less than 10 yrs.
		5,243	4. 10 yrs., less than 15 yrs.
		32,468	5. 15 years or more
		3,051	9. DK refused
		8,683	Blank. Not applicable (foreign born)
145	L-9c		YEARS LIVED IN UNITED STATES
		274	1. Less than 1 year
		1,339	2. 1 yr., less than 5 yrs.
		1,667	3. 5 yrs., less than 10 yrs.
		1,144	4. 10 yrs., less than 15 yrs.
		3,997	5. 15 years or more
		262	9. DK refused
		54,719	Blank. Not applicable (U.S. born)
146-171	-	-	BLANK

File Locations	Item No.	Frequency	Items and Codes
172-177	-	-	FINAL QUARTER BASIC WEIGHT BEFORE AGE-SEX-RACE/ETHNICITY ADJUSTMENT (has one implied decimal)
178			SAMPLING QUARTER
		16,902	1. Quarter 1
		19,814	2. Quarter 2
		13,470	3. Quarter 3
		13,216	4. Quarter 4
179-181	-	-	BLANK
182	Unit Control File		REGION
		12,500	1. Northeast
		13,606	2. Midwest
		21,982	3. South
		15,314	4. West
183	Unit Control File		GEOGRAPHIC DISTRIBUTION
			MSA Size
		8,007	1. 5,000,000 or more
		7,914	2. 2,500,000 - 4,999,999
		15,124	3. 1,000,000 - 2,499,999
		6,747	4. 500,000 - 999,999
		7,811	5. 250,000 - 499,999
		4,520	6. 100,000 - 249,999
		810	7. Under 100,000
		12,469	Blank. Non-MSA
184-185	-	-	BLANK
186	Unit Control File		MSA - NON-MSA RESIDENCE
		20,826	1. In MSA; in Central City
		30,107	2. In MSA; not in Central City
		12,469	3. Not in MSA
187-189	-	-	BLANK

File Locations	Item No.	Frequency	Items and Codes
190-200	-	-	CHRONIC CONDITION PREVALENCE AND INCIDENCE FACTOR (XX.XXXXXXXXX) - character format with implied decimal
			FINAL BASIC WEIGHT
201-209	-	-	QUARTER
210-218	-	-	SEMI-ANNUAL (QUARTER/2)
219-227	-	-	ANNUAL (QUARTER/4)

File Locations	Item No.	Frequency	Items and Codes
			6.5 WEIGHT
228-236	-	-	QUARTER, SEMI-ANNUAL AND ANNUAL*
			ESTIMATED RESTRICTED ACTIVITY DAYS IN PAST 2 WEEKS
237-245	-	-	QUARTER, SEMI-ANNUAL AND ANNUAL*
			ESTIMATED BED DAYS IN PAST 2 WEEKS
246-254	-	-	QUARTER, SEMI-ANNUAL AND ANNUAL*
			ESTIMATED WORK-LOSS DAYS IN PAST 2 WEEKS
255-263	-	-	QUARTER, SEMI-ANNUAL AND ANNUAL*
			ESTIMATED SCHOOL-LOSS DAYS IN PAST 2 WEEKS
264-272	-	-	QUARTER, SEMI-ANNUAL AND ANNUAL*
			ESTIMATED DOCTOR VISITS IN PAST 12 MONTHS
273-281	-	-	QUARTER
282-290	-	-	SEMI-ANNUAL
291-299	-	-	ANNUAL
			ESTIMATED SHORT-STAY HOSPITAL EPISODE DAYS IN PAST 12 MONTHS
300-308	-	-	QUARTER
309-317	-	-	SEMI-ANNUAL
318-326	-	-	ANNUAL
327-335	-	-	ANNUAL ESTIMATED NUMBER OF SHORT-STAY HOSPITAL EPISODES IN PAST 12 MONTHS

* Estimates can be made for these periods depending on whether one quarter, two quarters or all four quarters of records are used.

1996 NHIS ACCESS TO CARE PUBLIC USE FILE

File Locations	Item No.	Frequency	Items and Codes
336	-	-	BLANK
337-340	Recode	-	STRATUM FOR VARIANCE ESTIMATION
341	Recode	-	PSU FOR VARIANCE ESTIMATION
342-343	Recode	-	SUBSTRATUM FOR VARIANCE ESTIMATION
344-350	Generated	-	SECONDARY SAMPLING UNIT
351	Unit Control File		TYPE OF PSU
		41,421	1. Self representing
		21,981	2. Non self representing
352	Unit Control File		PANEL 4
		63,402	1-4. Code used to identify nationally representative subsamples
353		-	NSR STATUS VARIABLE
354-357		-	COLLAPSED VARIANCE STRATUM
358		-	VARIANCE PSU
359-399	-	-	BLANK

File Locations	Item No.	Frequency	Items and Codes
400	Recode		DUMMY RECORD FLAG
		2,586	1. Dummy record for locations 401-440
		60,816	Blank. Not a dummy record
401	1a		IS THERE USUAL PERSON/ PLACE FOR MEDICAL CARE
		52,922	1. Yes
		7,353	2. No
		186	3. More than one
		2,722	8. Not ascertained
		219	9. DK or refused
402	1b		ONE PLACE MOST OFTEN (Has more than one usual person or place Q 1a = 3)
		127	1. Yes
		47	2. No
		10	8. Not ascertained
		2	9. DK or refused
		63,216	Blank. NA
403-404	2		MAIN REASON NO USUAL SOURCE
		289	01. Two or more usual doctors/places
		3,377	02. Doesn't need doctor
		244	03. Doesn't like/trust/ believe in doctors
		140	04. Doesn't know where to go
		469	05. Previous doctor not available/moved
		1,689	06. No insurance/can't afford it
		6	07. Speaks different language
		32	08. No care available/ not convenient
		512	09. Changed residence
		327	10. Other reason
		261	98. Not ascertained
		66	99. DK or refused
		55,990	Blank. NA - Has/don't know if has one usual source of care

File Locations	Item No.	Frequency	Items and Codes
405	3a		ONE PLACE MOST OFTEN (Q 1a = 2 and Q 2 = 01)
		97	1. Yes
		74	2. No
		68	8. Not ascertained
		3	9. DK or refused
		63,160	Blank. NA
406	3b		PLACE FOR PREVENTIVE CARE (Q 1a = 2 and Q 2 = 01 and Q 3a NE 1)
		81	1. Yes
		84	2. No
		22	3. Not ascertained
		5	4. DK or refused
		63,210	Blank. NA
407	4a		GO ANY PLACE IN PAST 12 MONTHS FOR MEDICAL CARE
		2,728	1. Yes
		4,464	2. No
		2,912	8. Not ascertained
		152	9. DK or refused
		53,146	Blank. NA - Has a usual source of care
408-409	4b		KIND OF PLACE
		355	01. Hospital emergency room
		210	02. Urgent care/walk-in clinic
		1,118	03. Doctor's office
		479	04. Clinic
		164	05. Health center
		139	06. Hospital outpatient clinic
		84	07. HMO/Prepaid group
		39	08. Military or VA health care facility
		92	09. Some other place
		23	98. Not ascertained
		25	99. DK or refused
		60,674	Blank. NA

File Locations	Item No.	Frequency	Items and Codes
410	4c		WOULD STILL GO TO PLACE (Does not have a usual source of care, care but did go to a place for medical care sometime in the past 12 months, Q 4a = 1)
		1,462	1. Yes
		991	2. No
		23	8. Not ascertained
		252	9. DK or refused
		60,674	Blank. NA
411-412	4d		MAIN REASON WOULD NOT USE PLACE NOW (Does not have a usual source of care but did go to a place for medical care sometime in the past 12 months and would not use this place now Q 4c = 2)
		410	01. Changed residence/moved
		24	02. Changed jobs
		30	03. Employer changed insurance coverage
		64	04. Former usual source not available
		5	05. Owed money to former source
		76	06. Dissatisfied with former source/like new source better
		148	07. Medical care needs changed
		14	08. Former source stopped taking insurance/coverage
		199	09. Some other reason
		13	98. Not ascertained
		8	99. DK or refused
		62,411	Blank. NA

File Locations	Item No.	Frequency	Items and Codes
413	Recode		REGULAR SOURCE OF MEDICAL CARE
		52,922	1. Single regular source
		224	2. Multiple regular sources, but sees one most often
		204	3. Multiple regular sources, but no one source seen most often
		2,524	4. No current regular source but had a regular source sometime in the past year
		4,358	5. No current regular source of care and none in past year
		229	6. No current regular source and unknown if any in past year
		2,722	8. Not ascertained
		219	9. DK or refused
414-415	5a		KIND OF PLACE
		759	01. Hospital emergency room
		712	02. Urgent care/walk-in clinic
		35,675	03. Doctor's office
		7,430	04. Clinic
		1,815	05. Health center
		1,381	06. Hospital outpatient clinic
		3,713	07. HMO/prepaid group
		763	08. Military or VA health care facility
		268	09. Some other place
		516	98. Not ascertained
		114	99. DK or refused
		10,256	Blank. NA - Does not have/DK if have a usual source of care

File Locations	Item No.	Frequency	Items and Codes
416	5b		IS THERE PARTICULAR PERSON USUALLY SEEN
		44,452	1. Yes
		6,290	2. No
		587	8. Not ascertained
		346	9. DK or refused
		11,727	Blank. NA - Does not have/DK if have a usual source of care
417	6a		TYPE OF HEALTH PROFESSIONAL (Usually sees a particular person when goes to usual source of care. Q 5b = 1)
		43,371	1. Doctor
		79	2. Nurse
		222	3. Nurse practitioner
		180	4. Physician's assistant
		43	5. Chiropractor
		43	6. Some other professional
		500	8. Not ascertained
		14	9. DK or refused
		18,950	Blank. NA
418	6b		TYPE OF DOCTOR (Usually sees a particular doctor when goes to usual source of care Q 6a = 1)
		40,294	1. Family doctor/GP/ internist/pediatrician
		1,047	2. Obstetrician/gynecologist
		1,434	3. Other specialist
		451	8. Not ascertained
		145	9. DK or refused
		20,031	Blank. NA

File Locations	Item No.	Frequency	Items and Codes
419	7		LAST TIME WENT TO USUAL PLACE (Has a usual place for medical care)
		1,322	0. Hasn't been there yet/never
		20,540	1. Less than 3 months ago
		9,747	2. At least 3 months, less than 6 months ago
		9,693	3. At least 6 months, less than 1 year ago
		6,342	4. At least 1 year, less than 2 years ago
		4,098	5. Two or more years ago
		518	8. Not ascertained
		886	9. DK or refused
		10,256	Blank. NA
420	8		SAME PLACE FOR ROUTINE CARE (Has a usual place for medical care)
		50,176	1. Yes
		1,793	2. No
		914	8. Not ascertained
		263	9. DK or refused
		10,256	Blank. NA

File Locations	Item No.	Frequency	Items and Codes
421	Recode		RELATIONSHIP BETWEEN SOURCES OF SICK AND PREVENTIVE CARE
		50,176	1. Has the same usual source for sick and routine/preventive care
		1,793	2. Has a usual source for sick care, but it is not usual source of routine/routine/preventive care (may include persons with no usual source of routine/preventive care)
		1,177	3. Has a usual source of sick care, but unknown if has a usual source of routine/preventive care
		81	4. Has two or more usual sources of sick care and has a source for routine/preventive care
		84	5. Has two or more usual sources of sick care, but does not have a source for routine/ preventive care
		27	6. Has two or more usual sources of sick care, but unknown if has a usual source of routine/preventive care
		7,123	7. No usual source of sick care; not asked about preventive care
		2,722	8. Not ascertained
		219	9. DK or refused
422	9		GO TO OTHER PLACE PAST 12 MONTHS (Has a usual place for medical care)
		15,094	1. Yes
		36,270	2. No
		1,397	8. Not ascertained
		385	9. DK or refused
		10,256	Blank. NA

File Locations	Item No.	Frequency	Items and Codes
423	10b		CHANGED PLACE FOR ROUTINE CARE
		4,548	1. Yes
		55,862	2. No
		2,841	8. Not ascertained
		151	9. DK or refused
424-425	10d		MAIN REASON FOR CHANGE LAST TIME
		1,503	01. Changed residence
		197	02. Changed jobs
		643	03. Employer changed insurance coverage
		347	04. Former usual source not available
		4	05. Owed money to former source
		673	06. Dissatisfied with former source/like new source better
		308	07. Medical care needs changed
		126	08. Former source stopped taking insurance/coverage
		639	09. Some other reason
		103	98. Not ascertained
		5	99. DK or refused
		58,854	Blank. NA - Did not/DK change usual place of medical care in past 12 months
426	11b		NEEDED CARE BUT NOT ABLE TO GET IT
		1,703	1. Yes
		58,773	2. No
		2,792	8. Not ascertained
		134	9. DK or refused

File Locations	Item No.	Frequency	Items and Codes	
427-428	11d		MAIN REASON DIDN'T GET CARE	
		801	01.	Could not afford it
		278	02.	No insurance
		32	03.	Doctor did not accept Medicaid/ insurance plan
		50	04.	Insurance didn't cover
		30	05.	Not serious enough
		34	06.	Wait too long in clinic/office
		141	07.	Difficulty getting an appointment
		19	08.	Doesn't like/trust/ believe in doctors
		33	09.	No doctor available
		8	10.	Didn't know where to go
		14	11.	No way to get there
		7	12.	Hours not convenient
		0	13.	Speak a different language
		1	14.	Health of another family member interfered
		9	15.	Clinic/office not accessible
		135	16.	Other reason
		106	98.	Not ascertained
		5	99.	DK or refused
		61,699	Blank.	NA - Obtained medical care when needed during the past 12 months

File Locations	Item No.	Frequency	Items and Codes
429	11e		LACK OF INSURANCE OR MONEY A REASON
		121	1. Yes
		316	2. No
		102	3. Not ascertained
		3	9. DK or refused
		62,860	Blank. NA - Obtained medical care when needed during the 12 months or already mentioned that money or lack of insurance was a reason
430	12b		DELAYED SEEKING MEDICAL CARE BECAUSE OF COST
		4,507	1. Yes
		55,972	2. No
		2,812	8. Not ascertained
		111	9. DK or refused
431	13b		NEEDED DENTAL CARE BUT COULD NOT GET IT
		4,800	1. Yes
		55,603	2. No
		2,833	8. Not ascertained
		166	9. DK or refused
432	14b		NEEDED PRESCRIPTION MEDICINES BUT COULD NOT GET THEM
		1,504	1. Yes
		58,945	2. No
		2,781	8. Not ascertained
		172	9. DK or refused
433	15b		NEEDED EYEGLASSES BUT COULD NOT GET THEM
		1,874	1. Yes
		58,379	2. No
		2,947	8. Not ascertained
		202	9. DK or refused

File Locations	Item No.	Frequency	Items and Codes
434	16b		NEEDED MENTAL HEALTH CARE BUT COULD NOT GET IT
		348	1. Yes
		60,046	2. No
		2,849	8. Not ascertained
		159	9. DK or refused
435-436	-	-	BLANK
437-438	Check Item A5		CHECK ITEM (Reference Person Only) PERSON WHO ANSWERED MOST QUESTIONS
		865	00. Respondent not ascertained
		59,837	01-28. Person number
		0	30-97. Person number
		101	98. Active duty military
		13	99. Non household member
		2,586	Blank. Dummy record
439	Recode		RESPONDENT FOR SUPPLEMENT STATUS
		23,701	1. Self
		36,319	2. Proxy
		796	9. Unknown respondent
		2,586	Blank. Dummy record
440	-	-	BLANK

APPENDIX A

There is no Appendix for this document

APPENDIX B
INDUSTRY RECODES OUTLINE Revised in 1995

Recodes		Industry Title	SIC Code*
No. 1 Chrs. 80-81	No. 2 Chrs. 82-83		
01	01	AGRICULTURE	01-02,071-072,074-076,078
02	01	FORESTRY AND FISHERIES	08-09
10	02	MINING	10,12-14
20	03	CONSTRUCTION	15-17
(30-34, 40-46)	(04)	MANUFACTURING:	
(30-34)		NONDURABLE GOODS	
30	04	Food and kindred products	201-209
31	04	Textile mill and finished textile products	221-229,231-239
32	04	Printing, publishing and allied industries	271-279
33	04	Chemicals and allied products	281-287,289
34	04	Other nondurable goods	21,261-263,265,267,291,295,299,301-306,308,311,313-317,319

*Standard Industrial Classification

APPENDIX B
INDUSTRY RECODES OUTLINE Revised in 1995

Recodes		Industry Title	SIC Code*
No. 1 Chrs. 80-81	No. 2 Chrs. 82-83		
(30-34, 40-46)	(04)	MANUFACTURING: - continued	
(40-46)		DURABLE GOODS	
40	04	Furniture, lumber and wood	241-245,249,25
41	04	Primary metal industries	331-332,334,3331, 3334,3339,3351, 3353-3357,3363- 3366,3369,339
42	04	Fabricated metal industries, including ordnance	341-349
43	04	Machinery, except electrical	351-359
44	04	Electrical machinery, equipment and supplies	361-367,369
45	04	Transportation equipment	371-376,379
46	04	Other and not specified durable goods	321-329,381-382, 384-387,39

*Standard Industrial Classification

APPENDIX B
INDUSTRY RECODES OUTLINE Revised in 1995

Recodes		Industry Title	SIC Code*
No. 1 Chrs. 80-81	No. 2 Chrs. 82-83		
(50-54)	(05)	TRANSPORTATION, COMMUNICATIONS AND OTHER PUBLIC UTILITIES	
50	05	Railroads	40
51	05	Trucking service and warehousing	421-423
52	05	Other transportation	41,43-47
53	05	Communications	481-484,489
54	05	Utilities and sanitary	491-497
60	06	WHOLESALE TRADE	501-509,511-519

*Standard Industrial Classification

APPENDIX B
INDUSTRY RECODES OUTLINE Revised in 1995

Recodes		Industry Title	SIC Code*
No. 1 Chrs. 80-81	No. 2 Chrs. 82-83		
(61-65)	(07)	RETAIL TRADE	
61	07	General merchandise stores	531,533,539
62	07	Food, bakery and dairy stores	541-546,549
63	07	Automotive dealers and gasoline stations	551-557,559
64	07	Eating and drinking places	58
65	07	Other and not specified retail trade	521,523,525-527,56, 571-572,5731,5734-5736,591-594,5961-5963,598,5992-5995, 5999
(70-71)	(08)	FINANCE, INSURANCE, AND REAL ESTATE	
70	08	Banking and credit agencies	60-61
71	08	Insurance, real estate, and other finance	62-65,67

*Standard Industrial Classification

APPENDIX B
INDUSTRY RECODES OUTLINE Revised in 1995

Recodes		Industry Title	SIC Code*
No. 1 Chrs. 80-81	No. 2 Chrs. 82-83		
(75-85)	(09-12)	SERVICES:	
(75-76)	(09)	BUSINESS AND REPAIR SERVICES	
75	09	Business services	731-738,751,752, 7542
76	09	Repair services	753,7549,762-764, 7692,7694,7699
(77-78)	(10)	PERSONAL SERVICES	
77	10	Private households	88
78	10	Other personal services	701-704,721-726,729
79	11	ENTERTAINMENT AND RECREATION SERVICES	781-784,791-794,799
(80-85)	(12)	PROFESSIONAL AND RELATED SERVICES	
80	12	Hospitals	806
81	12	Health services, except hospitals	801-803,8041-8043, 8049,805,807-809
82	12	Elementary and secondary schools and colleges	821-822
83	12	Other educational services	823-824,829
84	12	Social services, religious and membership organizations	832-833,835-836, 839,84,861-866,869
85	12	Legal, engineering and other professional services	81,871-874,899

*Standard Industrial Classification

APPENDIX B
INDUSTRY RECODES OUTLINE Revised in 1995

Recodes		Industry Title	SIC Code*
No. 1 Chrs. 80-81	No. 2 Chrs. 82-83		
90	13	PUBLIC ADMINISTRATION	911-913,919,92-97
95	14	UNKNOWN INDUSTRY (Includes never worked)	-
96	14	REFUSED, CLASSIFIED, ETC.	
97	15	NOT IN LABOR FORCE - codes Blank and 8 in current activity recode (loc. 75) (Under 18 or 18+ and not in Labor Force).	
98	16	ARMED FORCES (excludes Reserves and National Guard)	

*Standard Industrial Classification

APPENDIX B
INDUSTRY RECODE TITLES

Code	Titles	Recode No. 1 Inclusions
01	AGRICULTURE, FORESTRY AND FISHERIES	01,02
02	MINING	10
03	CONSTRUCTION	20
04	MANUFACTURING	30-34, 40-46
05	TRANSPORTATION, COMMUNICATIONS AND OTHER PUBLIC UTILITIES	50-54
06	WHOLESALE TRADE	60
07	RETAIL TRADE	61-65
08	FINANCE, INSURANCE, AND REAL ESTATE	70-71
09	BUSINESS AND REPAIR SERVICES	75-76
10	PERSONAL SERVICES	77-78
11	ENTERTAINMENT AND RECREATION SERVICES	79
12	PROFESSIONAL AND RELATED SERVICES	80-85
13	PUBLIC ADMINISTRATION	90
14	UNKNOWN (includes never worked, refused, classified, etc.)	95-96
15	NOT IN LABOR FORCE	97
16	ARMED FORCES	98

APPENDIX C

OCCUPATION RECODE OUTLINE — Revised in 1995

Recodes			
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90	Occupation Title	SOC Code*
(01-03)	(01)	EXECUTIVE, ADMINISTRATIVE, AND MANAGERIAL OCCUPATIONS	-
01	01	Officials and administrators, public administration	111-113
02	01	Managers and administrators, except public administration	121-128,131-1344,1351-1354,1359, 136-139
03	01	Management related occupations	1412,1414-1415, 1419,142-143, 1442-1443,1449, 145,1472-1473,149

*Standard Occupational Classification

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
(04-11)	(02)	PROFESSIONAL SPECIALTY OCCUPATIONS	-
04	02	Engineers	1622-1628,1632-1637,1639
05	02	Architects and surveyors	161,164
06	02	Natural mathematical and computer scientists	171-172,1732-1733,1739, 1842-1843, 1845-1847,1849, 1852-1855
07	02	Health diagnosing occupations	261-262,27,281, 283,289
08	02	Health assessment and treating occupations	29,301-302, 3031-3034,3039, 304
09	02	Teachers, librarians and Counselors	2212-2218, 2222-2228, 2231-2238, 2242-2247, 2249,231-233, 235,236,239,24, 251,252
10	02	Writers, artists, entertainers and athletes	34,321-329, 331-333,398
11	02	Other professional specialty occupations	1912-1916, 1919,192, 2032-2033, 2042,2049, 211-212

*Standard Occupational Classification.

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
(12-13)	(03)	TECHNICIANS AND RELATED SUPPORT OCCUPATIONS	-
12	03	Health technologists and technicians	362-366,369
13	03	Technologists, technicians except health	3711-3713,3719, 372-373,382, 3831-3833, 384,389, 392-393,396, 3971-3972, 3974,399, 825
(14-16)	(04)	SALES OCCUPATIONS	-
14	04	Supervisors and proprietors	40
15	04	Sales representatives, commodities and finance	4122-4124, 4152-4153, 421,423-424
16	04	Other sales	4342-4348,4351-4354,4356,4359, 4362-4367,4369, 444-447,449

*Standard Occupational Classification

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
(17-21)	(05)	ADMINISTRATIVE SUPPORT OCCUPATIONS, INCLUDING CLERICAL	-
17	05	Computer equipment operators	4612-4613
18	05	Secretaries, stenographers and typists	4622-4624
19	05	Financial records processing occupations	4712-4713, 4715-4716, 4718
20	05	Mail and message distributing	4742-4745
21	05	Other administrative support	4511-4514,4516, 4519,4521-4529, 463,4642-4645, 4649,4662-4664, 4692,4694,4696, 4699,4722-4723, 4729,4732-4733, 4739,4751-4759, 4782-4784,4786- 4787,4791-4795, 4799
22	06	Private household occupations	502-507,509

*Standard Occupational Classification.

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
(23-24)	(07)	PROTECTIVE SERVICE OCCUPATIONS	
23	07	Police and firefighters	5111-5112, 5122-5123, 5132-5134
24	07	Other protective service occupations	5113,5142,5144, 5149
(25-28)	(08)	SERVICE OCCUPATIONS, EXCEPT PROTECTIVE AND HOUSEHOLD	
25	08	Food service	5211-5219
26	08	Health service	5232-5233,5236
27	08	Cleaning and building service	5241-5242, 5244-5246,5249
28	08	Personal service	5251-5258,5262-5264,5269
(29-31)	(09)	FARMING, FORESTRY AND FISHING OCCUPATIONS	
29	09	Farm operators and managers	5512-5515,5522-5525
30	09	Farm workers and other agricultural workers	5611-5619, 5621-5622, 5624-5625,5627
31	09	Forestry and fishing occupations	571-573,579, 583-584,8241(pt.)

*Standard Occupational Classification.

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
(32-34)	(10)	PRECISION PRODUCTION, CRAFT AND REPAIR OCCUPATIONS	
32	10	Mechanics and repairers	60,6111-6118, 613-614,6151-6159,616,6171-6179
33	10	Construction and extractive trades	6311-6316,6318, 632,6412-6414(pt.),6422, 6424,6432-6433, 6442-6444,645, 6462-6468,6472-6476,6479,652-654,656
34	10	Precision production occupations	67,71, 6811-6814, 6816-6817, 6821-6824, 6829,6831-6832, 6835,6839,6844, 6852-6854,6856, 6859,6861-6862, 6864-6867,6869, 6871-6873,6879, 6881-6882,691-696,7477(pt.), 7668,7677(pt.), 7752,828

*Standard Occupational Classification

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
		OPERATORS, FABRICATORS AND LABORERS	
(35-36)	(11)	MACHINE OPERATORS, ASSEMBLERS AND INSPECTORS	
35	11	Machine operators and tenderers, except precision	6841-6842,6849, 6855,6863,6868, 7312-7319,7322, 7324,7326,7329, 7339,7342-7344, 7349,7431-7435, 7439,7443-7444, 7449,7451-7452, 7459,7462-7463, 7467,7472,7474, 7476-7478,7479, 7512-7519,7522, 7529,7539,7542- 7544,7549,7631- 7636,7639,7642- 7644,7649,7651- 7652,7654-7659, 7661-7667,7669, 7671-7676, 7677(pt.), 7678-7679
36	11	Fabricators, assemblers, inspectors and samplers	7332-7333,7532- 7533,7714,7717, 72,774,7753- 7759,782-785,787

*Standard Occupational Classification

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
(37-39)	(12)	TRANSPORTATION AND MATERIAL MOVING OCCUPATIONS	
37	12	Motor vehicle operators	8111,8212-8216, 8218-8219,874
38	12	Other transportation, except motor vehicles	8113,8232-8233, 8239,8241(pt.), 8242-8245
39	12	Material moving equipment operators	812,8312-8319
(40-41)	(13)	HANDLERS, EQUIPMENT CLEANERS, HELPERS AND LABORERS	
40	13	Construction laborers	871
41	13	Freight, stock and material handlers	85,861-863, 8641-8646, 8648,865, 8722-8726,873, 875,8761,8769

* Standard Occupational Classification

APPENDIX C

OCCUPATION RECODE OUTLINE Revised in 1995

Recodes		Occupation Title	SOC Code*
No. 1 Chrs. 87-88	No. 2 Chrs. 89-90		
95	14	UNKNOWN OCCUPATION (Includes never worked)	
96	14	REFUSED, CLASSIFIED, ETC.	
97	15	NOT IN LABOR FORCE - codes Blank and 8 in current activity recode (Loc. 75). (Under 18 or 18+ and Not in Labor Force)	
98	16	MILITARY	

*Standard Occupational Classification

APPENDIX C
OCCUPATION RECODE TITLES

Code	Titles	Recode No. 1 Inclusions
	MANAGERIAL AND PROFESSIONAL SPECIALTY OCCUPATIONS	
01	EXECUTIVE, ADMINISTRATIVE AND MANAGERIAL OCCUPATIONS	01-03
02	PROFESSIONAL SPECIALTY OCCUPATIONS	04-11
	TECHNICAL, SALES AND ADMINISTRATIVE SUPPORT OCCUPATIONS	
03	TECHNICIANS AND RELATED SUPPORT OCCUPATIONS	12-13
04	SALES OCCUPATIONS	14-16
05	ADMINISTRATIVE SUPPORT OCCUPATIONS, INCLUDING CLERICAL	17-21
	SERVICE OCCUPATIONS	
06	PRIVATE HOUSEHOLD OCCUPATIONS	22
07	PROTECTIVE SERVICE OCCUPATIONS	23-24
08	SERVICE OCCUPATIONS, EXCEPT PROTECTIVE AND HOUSEHOLD	25-28
09	FARMING, FORESTRY AND FISHING OCCUPATIONS	29-31
10	PRECISION PRODUCTION, CRAFT AND REPAIR OCCUPATIONS	32-34
	OPERATORS, FABRICATORS AND LABORERS	
11	MACHINE OPERATORS, ASSEMBLERS AND INSPECTORS	35-36
12	TRANSPORTATION AND MATERIAL MOVING OCCUPATIONS	37-39
13	HANDLERS, EQUIPMENT CLEANERS, HELPERS AND LABORERS	40-41
14	UNKNOWN OCCUPATION (includes never worked, refused, classified, etc.)	95-96
15	NOT IN LABOR FORCE	97
16	MILITARY	98

VARIANCE ESTIMATION FOR PERSON DATA USING THE NHIS PUBLIC USE PERSON DATA TAPE, 1995-2004

April 17, 1998

About this document:

This document provides basic design information about the 1995-2004 NHIS and presents methods to compute standard errors for each annually released person-level database. This document focuses upon a full-sample NHIS Core survey that is anticipated for each data collection year. For some years the full-sample methods need to be modified to account for design changes. In particular, the 1996 NHIS has a sample design quite different from the 1995 NHIS. Also, Supplemental surveys may require modified methods. Some notes about these modifications appear at the end of this document.

Contents

VARIANCE ESTIMATION FOR PERSON DATA USING THE NHIS PUBLIC USE PERSON DATA TAPE, 1995

Introduction: The data collected in the NHIS are obtained through a complex sample design involving stratification, clustering, and multistage sampling, and the final weights are subject to several adjustments. Any variance estimation methodology must involve numerous simplifying assumptions about the design and weighting. We provide some oversimplified conceptual NHIS design structures that should allow users of this Public Use Data Set to compute reasonably accurate standard errors.

There are several available software packages for analyzing complex samples. A comparison is beyond the scope of this document, but an Internet web page Summary of Survey Analysis Software currently located at http://www.fas.harvard.edu/~stats/survey-soft/survey-soft.html provides references and discussion. At NCHS the software package SUDAAN has been used to produce standard errors. In this document SAS and SUDAAN computer code is provided, but without guarantees of any kind. The computer code and methods are subject to change without notification to the user. The entire risk as to the results and performance is assumed by the user. NCHS recommends that any analysis of NHIS data be done under the supervision of a statistician who understands the implications of complex-sample design surveys.

Conceptual NHIS design for 1995 The U.S. Bureau of the Census partitions the state counties or equivalents along with metropolitan areas into a universe of about 1900 Primary Sampling Units (PSUs) (note, PSUs may be combined counties) to provide the primary sampling areas for its many national surveys. For the NHIS these universe PSUs are partitioned into geographical strata at the state level. Some of the larger universe PSUs are self-representing (SR), i.e., they are in the NHIS with certainty. The other PSUs are called non-self-representing (NSR) or non-certainty PSUs. Within each state the NSR PSUs are partitioned into strata based upon similarity of PSU characteristics. Within each NSR stratum 2 PSUs are selected using Durbin's probability proportional to size (PPS) sampling method using the population as a measure of size. (In some smaller states only 1 PSU is drawn PPS). The SR PSUs are equivalent to strata, but historically they have been referred to as PSUs. (PPS and Durbin sampling are discussed in Chapter 9A of Cochran (1977)).

Within a sampled NSR or SR PSU the geography is partitioned into smaller geographical clusters which are used to form the universe of secondary sampling units (SSUs). These SSUs are then partitioned into density strata based upon black and Hispanic population concentration as determined by the 1990 Decennial Census. An additional strata for new construction since the last Decennial Census is also created. Within each density stratum SSUs are sampled at different rates to meet different design objectives. Within each sample SSU, all households containing black or Hispanic persons are sampled, while all other households are subsampled. Supplemental NHIS surveys may require additional sampling at SSU, household, or family levels.

The fundamental sampling weights are created such that under ideal sampling conditions, unbiased estimators for each level of sampling are available. In practice, however, the final sampling weights are adjusted for non-response, and ratio adjusted. Furthermore, in 1995 a government shutdown resulted in three lost weeks of sample which resulted in further weighting adjustments. The most important adjustment is a quarterly post-stratification to 90 age/sex/race/ethnicity Census control totals.

For variance estimation purposes, NCHS treats the NHIS as a two-stage sample. The PSU probabilities of selection are known, and the SSUs are treated as sampled with replacement within PSU density strata. Sampling weights are adjusted by postratification. With these assumptions the SUDAAN software is used to compute variances. Much of the design information, state, density strata, and Durbin probabilities can be used to identify the smaller geographical areas. NCHS forbids the disclosure of information which may compromise the confidentiality promised to survey respondents, so some design information is not provided with the Public Use Data. While all design information is not available to the public, variance estimation methods exist which provide similar results to the NCHS internally used methodology. Two methods are described below.

Design Information Available on the NHIS Public Use Databases.

CAUTION For 1996 databases, refer to the Notes at the end of this document.

The following variables are used to produce code for variance estimation. Field locations below are from the PERSON level database, but may change on other databases; the user should check the file documentation.

Variable Name	Tape Location	Field Label
STRAT_V	337-340	'STRATA FOR VARIANCE ESTIMATION'
PSU_V	341	'PSU FOR VARIANCE ESTIMATION'
SUB_V	342-343	'SUBSTRATUM FOR VARIANCE ESTIMATION'
SSU	344-350	'SECONDARY SAMPLING UNIT'
PANEL	352	'PANEL 4'
TYPE_PSU	351	'TYPE OF PSU'
WTF	219-227	'FINAL BASIC WEIGHT'

Two methods of variance estimation are provided.

Method 1 - 187 Strata containing 2 PSUs per stratum sampled with replacement

Here, the NHIS universe has been partitioned into 187 strata. Most of the original NHIS strata and PSUs retain their original sampling structure with two PSUs being sampled per stratum, but a few strata have been collapsed, and in the largest self-representing strata, two pseudo-PSUs have been created. All PSUs are treated as sampled with replacement within their respective strata. This method will provide somewhat conservative standard errors, and the standard error estimator itself has less stability than the standard error estimator described by Method 2 below. Method 1 should be applicable to many complex survey sample design computer programs which require exactly 2 sampled PSUs per stratum. This method is robust when analyzing subsetted data (See the section "Subsetted Data Analysis" below).

Coding required (SAS code provided):

```
STRATUM  = STRAT_V;

PSU = PANEL;

IF (PSU_V = 5) THEN PSU = INT((PANEL + 1)/2);

IF(PSU_V = 8) THEN STRATUM  = 553;

IF((TYPE_PSU = 1) AND (PSU_V IN (2,4))) THEN STRATUM =
(STRAT_V -1);

IF((STRAT_V  = 921)  AND (PSU_V  = 3)) THEN STRATUM = 901;
```

As a check the user should observe 374 PSUs when using the full database.

For the above simplification of the NHIS sample design structure, the following SUDAAN design statements may be used. (Note, the input file must first be sorted by STRATUM and PSU variables.)

```
PROC ...  DESIGN = WR;
NEST      STRATUM PSU ;
WEIGHT    WTF;
```

See the Section "Worked SUDAAN Examples" below for further discussion.

Method 2 - Multiple PSUs per Stratum design sampled with replacement

This method provides for more statistically efficient variance estimation than Method 1, since it makes better use of the sampling design information. Its application is limited to software that can handle multiple PSUs per stratum, e.g., SUDAAN. For this method the original certainty PSUs are partitioned by aggregations of the original race-ethnic density strata used in sampling. The first randomly sampled unit is actually the SSU variable which is now treated as the PSU variable. (Note, a certainty PSU unit contributes nothing to the variance at the PSU sampling level.) Non-certainty-strata PSUs are treated as being sampled with replacement within their respective strata. Except for a few special cases, the non-certainty PSUs have exactly the same structure in both Methods 1 and 2.

Coding required, (SAS code provided):

```
IF TYPE_PSU = 1 THEN DO;  /* certainty strata PSUs  */

      STRATUM = STRAT_V*1000 + SUB_V;
         PSU  = SSU  ;
         END;

 ELSE DO ;                /* non-certainty PSU   */;

      STRATUM = STRAT_V ;
          PSU  =   PSU_V ;
         END;
```

As a check, the user should observe the following counts:

Certainty Strata PSUs	4079
Non-certainty Strata PSUs	259
Total PSUs	4338

For the Method 2 design structure, the following SUDAAN design statements may be used. (Note, the input file must first be sorted by STRATUM and PSU variables.)

```
PROC ...  DESIGN = WR;
NEST      STRATUM PSU;
WEIGHT    WTF;
```

See the Section "Worked SUDAAN Examples" for further discussion.

CAUTION. Method 2 should only be used on a full sample person data base. Using this method with subsetted data may lead to incorrectly computed standard errors. (See the section "Subsetted Data Analysis" below). If using a subsetted data set, the user should check the degree of agreement of the certainty and non-certainty counts with the values presented above.

CAUTION
A typically used rule-of-thumb for degrees of freedom to associate with a standard error is the quantity (number of PSUs - number of strata). This rule assumes that the PSUs are somewhat comparable in size. For Method 2 this rule may be grossly inaccurate since the concept of PSU is quite different for certainty and non-certainty strata. Certainty strata PSUs of Method 2 have small weighted values relative to those of non-certainty PSUs. The rule-of-thumb degrees of freedom for Method 1 is 187, and Method 2 should have a "true" degrees of freedom exceeding that of Method 1. However, for practical purposes, any degrees of freedom exceeding 120 can be treated as infinite, i.e., one uses a normal Z-statistic instead of a t-statistic for testing. Note, that a one-tailed critical t0.025 at 120 degrees of freedom is 1.98 while at an infinite degrees of freedom (i.e., a z-value) is 1.96. If a variable of interest covers most of the NHIS PSUs, the limiting value would probably be adequate for analysis. The user should consult a mathematical statistician for discussion of degrees of freedom.

SUBSETTED DATA ANALYSES

Frequently, studies of NHIS variables are restricted to select subdomains, e.g., persons aged 65 and older. To save on storage the user may delete all records outside of the domain of interest. This procedure of keeping only select records is called subsetting the data. With a subsetted data set one can produce correct point estimates, e.g., the subdomain means, but standard errors may be computed incorrectly when using a compromised design structure. For example, if a stratum of Method 2 contains 10 PSUs and 5 are lost because of subsetting, a SUDAAN run on the subsetted data will use an incorrect formula to compute stratum contributions to the variance. If the full data are run, SUDAAN correctly handles the 5 empty PSUs. Note, that SUDAAN has a SUBPOPN option that allows the targeting of a subdomain from a full design data base. (See the SUDAAN manual for details).

Subsetting methods with SUDAAN

Strategy 1. Use Method 1 above with the MISSUNIT option on the NEST statement -

```
        NEST        STRATUM PSU/MISSUNIT ;
```

If a WR design has exactly 2 PSUs per stratum and some PSUs are removed from the database then the SUDAAN MISSUNIT option performs a fix-up which produces a standard error identical to that achieved when using a full data set and SUBPOPN statement. Note, other output like design effects, degrees of freedom, standardization may be computed differently. The user is responsible for checking that subsetted input leads to correct results.

Strategy 2. Use Method 1 or 2 above on a "fixed-up" subsetted data set. Basically, one needs to add some dummy records containing full design information to the subsetted data set. To do this follow these instructions:

1. Create a 2-variable file containing STRATUM and PSU for each record of the full person file (100,000+ records)

2. Sort this file by STRATUM and PSU within STRATUM.

3. Keep only 1 record for each PSU
   add WTF = 10 -10 as a very small weight
   add variable DUMMY = 0 to designate dummy record

   A file, called DESIGN containing 4 variables with
    374 records (Method 1 used) or with
   4338 records (Method 2 used) is created

4. Append DESIGN to the original subsetted database, called DATASET, to form a new set, called DATANEW.

   Define DUMMY = 1 on the DATASET component.

   On the DESIGN component records define all variables other than STRATUM, PSU, WTF, DUMMY as missing ".".

5. Sort DATANEW by STRATUM PSU

6. In SUDAAN use a "SUBPOPN DUMMY = 1;" line to direct SUDAAN to restrict estimation to the subdomain of interest.

With the above fix-up SUDAAN will correctly handle empty PSUs when computing the standard errors. SUDAAN output that needs the entire full sample database for correct computation, e.g., design effects, may or may not be appropriate. See the SUDAAN manual for computational forms or consult with a mathematical statistician for correct interpretation.

Other notes on Subsetting data:

If a subsetted database under Method 2 has only a few missing PSUs, the subsetted database can probably be run with SUDAAN without being fixed up. For example, a subsetting by SEX will most likely result in all PSUs still being in sample, but black males aged 65 and older would result in the loss of many PSUs. The impact of running SUDAAN on uncorrected subsetted data varies. Frequently, subsetted runs produce results consistent with those run on a full data set, but sometimes they do not.

Subsetting by aggregates of Strata does not need a fix-up.

The condition, doctor visit, and hospital record databases are actually subsetted files. To use with SUDAAN properly, the information should be linked back to the appropriate person on the person file. Some statistics, based upon aggregation of records, may be computed directly from this file along with the fix-up. Consult with a statistician for appropriate SUDAAN usage.

WORKED SUDAAN EXAMPLES

In the following runs the variables used are

LDR = proportion of persons without a doctor visit in the last 2 years

TDV_R = mean number of annual doctor visits (based upon 2 week recall)

HLT_FP = proportion of persons with self-reported fair or poor health status (omitting missing)

AGE2: 1 = aged less than 18
2 = aged 18 to 44
3 = aged 45 to 64
4 = aged 65 and older

The following SUDAAN code was executed for both Method 1 and Method 2:

Caution The output presented below is based upon a preliminary NHIS Public Use database. Your Public Use database may produce slightly different SUDAAN output.

```
   PROC DESCRIPT DATA = HIS.infile  FILETYPE=SAS DESIGN = WR;

   NEST       STRATUM PSU ;
   WEIGHT       WTF;

   VAR     LDR    TDV_R     HLT_FP ;

   SUBGROUP     SEX  AGE2;
   LEVELS       2    4;
   TABLES     SEX   AGE2;

   PRINT    NSUM WSUM MEAN SEMEAN
            / WSUMFMT=F10.0     MEANFMT=F8.5    SEMEANFMT=F8.5   ;
```

Method 1: partial output:

```
                                S U D A A N
            Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute      April 1996
                                Release 7.00

Number of observations read    : 102467    Weighted count :261889548
Number of observations skipped :      0
(WEIGHT variable nonpositive)
Denominator degrees of freedom :    187

                              Research Triangle Institute

                                The DESCRIPT Procedure

by: Variable, SEX.
```

Variable		SEX Total	1	2
LDR	Sample Size	102467	48809	53658
	Weighted Size	261889549	127570237	134319312
	Mean	0.13797	0.18013	0.09793
	SE Mean	0.00178	0.00250	0.00178
TDV_R	Sample Size	102467	48809	53658
	Weighted Size	261889549	127570237	134319312
	Mean	5.90759	4.90385	6.86089
	SE Mean	0.09060	0.10039	0.12407
HLT_FP	Sample Size	101277	48266	53011
	Weighted Size	258963568	126221708	132741859
	Mean	0.10126	0.09124	0.11079
	SE Mean	0.00157	0.00188	0.00176

Method 1: partial output:

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute April 1996
Release 7.00

Number of observations read : 102467 Weighted count :261889548
Number of observations skipped : 0
(WEIGHT variable nonpositive)
Denominator degrees of freedom : 187

Research Triangle Institute

The DESCRIPT Procedure

by: Variable, AGE2.

Variable		AGE2 Total	1	2
LDR	Sample Size	102467	29711	40801
	Weighted Size	261889549	70670755	108040689
	Mean	0.13797	0.08894	0.18489
	SE Mean	0.00178	0.00269	0.00268
TDV_R	Sample Size	102467	29711	40801
	Weighted Size	261889549	70670755	108040689
	Mean	5.90759	4.29682	4.88589
	SE Mean	0.09060	0.09797	0.12432
HLT_FP	Sample Size	101277	29183	40423
	Weighted Size	258963568	69438212	107054300
	Mean	0.10126	0.02552	0.06610
	SE Mean	0.00157	0.00129	0.00168

Variable		3	4
LDR	Sample Size	20000	11955
	Weighted Size	51713265	31464840
	Mean	0.14461	0.07606
	SE Mean	0.00293	0.00251
TDV_R	Sample Size	20000	11955
	Weighted Size	51713265	31464840
	Mean	7.08504	11.09843
	SE Mean	0.17859	0.30642
HLT_FP	Sample Size	19834	11837
	Weighted Size	51315866	31155190
	Mean	0.16651	0.28344
	SE Mean	0.00356	0.00519

Method 2 Partial Output

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute April 1996
Release 7.00

Number of observations read : 102467 Weighted count :261889548
Number of observations skipped : 0
(WEIGHT variable nonpositive)
Denominator degrees of freedom : 4030

Research Triangle Institute

The DESCRIPT Procedure

by: Variable, SEX.

Variable		SEX Total	1	2
LDR	Sample Size	102467	48809	53658
	Weighted Size	261889549	127570237	134319312
	Mean	0.13797	0.18013	0.09793
	SE Mean	0.00174	0.00231	0.00184
TDV_R	Sample Size	102467	48809	53658
	Weighted Size	261889549	127570237	134319312
	Mean	5.90759	4.90385	6.86089
	SE Mean	0.07704	0.08503	0.11403
HLT_FP	Sample Size	101277	48266	53011
	Weighted Size	258963568	126221708	132741859
	Mean	0.10126	0.09124	0.11079
	SE Mean	0.00152	0.00174	0.00182

Method 2 Partial Output

S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute April 1996
Release 7.00

Number of observations read : 102467 Weighted count :261889548
Number of observations skipped : 0
(WEIGHT variable nonpositive)
Denominator degrees of freedom : 4030

Research Triangle Institute

The DESCRIPT Procedure

by: Variable, AGE2.

Variable		AGE2 Total	1	2
LDR	Sample Size	102467	29711	40801
	Weighted Size	261889549	70670755	108040689
	Mean	0.13797	0.08894	0.18489
	SE Mean	0.00174	0.00271	0.00254
TDV_R	Sample Size	102467	29711	40801
	Weighted Size	261889549	70670755	108040689
	Mean	5.90759	4.29682	4.88589
	SE Mean	0.07704	0.09116	0.11805
HLT_FP	Sample Size	101277	29183	40423
	Weighted Size	258963568	69438212	107054300
	Mean	0.10126	0.02552	0.06610
	SE Mean	0.00152	0.00118	0.00157

Variable		3	4
LDR	Sample Size	20000	11955
	Weighted Size	51713265	31464840
	Mean	0.14461	0.07606
	SE Mean	0.00303	0.00269
TDV_R	Sample Size	20000	11955
	Weighted Size	51713265	31464840
	Mean	7.08504	11.09843
	SE Mean	0.16109	0.28387
HLT_FP	Sample Size	19834	11837
	Weighted Size	51315866	31155190
	Mean	0.16651	0.28344
	SE Mean	0.00351	0.00501

Best NHIS design using Durbin probabilities (not available to the public)
and weights adjusted by post-stratification

Variable		SEX		
		Total	1	2
LDR	Sample Size	102467	48809	53658
	Weighted Size	261889549	127570237	134319312
	Mean	0.13784	0.17991	0.09789
	SE Mean	0.00170	0.00221	0.00182
TDV_R	Sample Size	102467	48809	53658
	Weighted Size	261889549	127570237	134319312
	Mean	5.90468	4.89733	6.86141
	SE Mean	0.07511	0.08320	0.11217
HLT_FP	Sample Size	101277	48266	53011
	Weighted Size	258974266	126232939	132741328
	Mean	0.10127	0.09125	0.11080
	SE Mean	0.00137	0.00159	0.00165

Best NHIS design using Durbin probabilities (not available to the public)
and weights adjusted by post-stratification

Post-stratified estimates

by: Variable, AGE2.

Variable		AGE2 Total	1	2
LDR	Sample Size	102467	29711	40801
	Weighted Size	261889549	70670755	108040689
	Mean	0.13784	0.08845	0.18484
	SE Mean	0.00170	0.00258	0.00248
TDV_R	Sample Size	102467	29711	40801
	Weighted Size	261889549	70670755	108040689
	Mean	5.90468	4.29787	4.87876
	SE Mean	0.07511	0.09066	0.11858
HLT_FP	Sample Size	101277	29183	40423
	Weighted Size	258974266	69441900	107059972
	Mean	0.10127	0.02555	0.06624
	SE Mean	0.00137	0.00116	0.00153

Variable		3	4
LDR	Sample Size	20000	11955
	Weighted Size	51713265	31464840
	Mean	0.14484	0.07587
	SE Mean	0.00298	0.00268
TDV_R	Sample Size	20000	11955
	Weighted Size	51713265	31464840
	Mean	7.08472	11.09687
	SE Mean	0.16180	0.27613
HLT_FP	Sample Size	19834	11837
	Weighted Size	51315313	31157082
	Mean	0.16633	0.28322
	SE Mean	0.00342	0.00487

Remark on Examples

A comparison of the three SUDAAN examples shows that Method 2 performs quite well when compared to the "best" internal NCHS variance design for the NHIS. Based on limited preliminary evidence, it appears that for means, Method 2 typically provides standard errors in close agreement with, while slightly larger than, the standard errors produced by the NCHS "best" method. Method 1 tends to provide slightly larger standard errors than Method 2 does, although the sample output does include examples where the Method 1 standard error is smaller than the Method 2 standard error.

Reference:

(1977) Cochran, W. G., Sampling techniques (3rd ed), John Wiley & Sons

Notes for Year 2000 application (added 01/21/98)

The variance estimation methods of this document may be applied to the Year 2000 Objectives Public Use File. The following changes must be made:

The design information variables are all in the same file locations with the exception of "WTF".
Substitute:

```
   WTF         207-212   'FINAL BASIC WEIGHT'
```

The PSU check for method 2 should now read:

As a check, the user should observe the following counts:

Certainty Strata PSUs	3804
Non-certainty Strata PSUs	259
Total PSUs	4063

Notes on the 1996 NHIS (added 04/17/98)

In 1996 the NHIS survey underwent a transition from a paper-and-pencil to a computer-assisted interview process. This transition resulted in roughly 5/8 of the available full sample being targeted for processing and public release. In 1997 the full sample was again implemented. For 1996 the reader should substitute the information on pages 3 and 4 and the top of page 5:

Design Information Available on the NHIS Public Use Databases.

Method 1 - 187 Strata containing 2 PSUs per stratum sampled with replacement

Method 2 - Multiple PSUs per Stratum design sampled with replacement

with the 1996 information on the following pages:

Design Information Available on the 1996 NHIS Public Use Databases.

The following variables are used to produce code for variance estimation. Field locations below are from the PERSON level database, but may change on other databases; the user should check the file documentation.

Variable Name	Location	Field Label
STRAT96*	354-357	'COLLAPSED VARIANCE STRATUM'
PSU96*	358	'VARIANCE PSU'
SUB_V	342-343	'SUBSTRATUM FOR VARIANCE ESTIMATION'
SSU	344-350	'SECONDARY SAMPLING UNIT'
PANEL	352	'PANEL 4'
NSR96*	353	'NSR STATUS VARIABLE'
WTF	219-227	'FINAL BASIC WEIGHT'

(*indicates modified design variables added to the 1996 databases)

Two methods of variance estimation are now provided.

Method 1.96 -98 Strata containing 3 PSUs per stratum sampled with replacement

Here, the NHIS universe has been partitioned into 98 collapsed strata with 3 PSUs per stratum. All PSUs are treated as sampled with replacement within their respective strata. This method will provide somewhat conservative standard errors, and this standard error estimator itself has less stability than the standard error estimator described by Method 2.96 below.

Coding required, (SAS code provided):

```
STRATUM = INT(STRAT96/10) * 10  ;

PSU = PANEL ;
```

Note, INT ( ) is the Integer-value SAS function, e.g., INT(2.3) = 2

As a check the user should observe 98*3 = 294 PSUs when using the full database.

For the above simplification of the NHIS sample design structure, the following SUDAAN design statements may be used. (Note, the input file must first be sorted by STRATUM and PSU variables).

```
PROC ... DESIGN = WR;
NEST     STRATUM PSU ;
WEIGHT   WTF;
```

Method 2.96 - Multiple PSUs per Stratum design sampled with replacement

This method provides for more statistically efficient variance estimation than Method 1.96, since it makes better use of the sampling design information. Its application is limited to software that can handle multiple PSUs per stratum, e.g., SUDAAN. For this method the original certainty PSUs are partitioned by aggregations of the original race-ethnic density strata used in sampling. The first randomly sampled unit is actually the SSU variable which is now treated as the PSU variable. (Note, a certainty PSU unit contributes nothing to the variance at the PSU sampling level). Non-certainty strata PSUs are treated as being sampled with replacement within their respective strata.

Coding required (SAS code provided):

```
IF NSR96 = 1 THEN DO;  /*1996 certainty strata PSUs  */

     STRATUM  = STRAT96*100 + SUB_V;
        PSU    = SSU
        END;

ELSE DO;              /* 1996 non-certainty PSU     */

   STRATUM = STRAT96 ;
     PSU     = PSU96   ;
     END;
```

As a check, the user should observe the following counts:

Certainty Strata PSUs	1736
Non-certainty Strata PSUs	240
Total PSUs	1976

For the Method 2.96 design structure, the following SUDAAN design statements may be used. (Note, the input file must first be sorted by STRATUM and PSU variables.)

```
    PROC ...  DESIGN = WR;
    NEST      STRATUM PSU ;
    WEIGHT    WTF ;
```

Caution. Both Method 1.96 and Method 2.96 should only be used on a full sample person database. Using this method with subsetted data may lead to incorrectly computed standard errors. (See the section Subsetted Data Analysis in the 1995 section). If using a subsetted data set, the user should check the degree of agreement in the PSU counts with the values presented above for either of the two methods. Unlike Method 1 for 1995, Method 1.96 is not robust for analyzing subsetted survey data.

CAUTION

A typically used rule-of-thumb for degrees of freedom to associate with a standard error is the quantity (number of PSUs - number of strata). This rule assumes that the PSUs are somewhat comparable in size. For Method 2.96 this rule may be grossly inaccurate since the concept of PSU is quite different for certainty and non-certainty strata. Certainty strata PSUs of Method 2.96 have small weighted values relative to those of non-certainty PSUs. The rule-of-thumb degrees of freedom for Method 1.96 is 196, and Method 2.96 should have a true degrees of freedom exceeding that of Method 1.96. However, for practical purposes, any degrees of freedom exceeding 120 can be treated as infinite, i.e., one uses a normal Z-statistic instead of a t-statistic for testing. Note, that a one-tailed critical t0.025 at 120 degrees of freedom is 1.98 while at an infinite degrees of freedom (i.e., a z-value) is 1.96. If a variable of interest covers most of the NHIS PSUs, the limiting value would probably be adequate for analysis. The user should consult a mathematical statistican for discussion of degrees of freedom.

The observant reader may notice that the 1996 method 1.96 has a larger rule of thumb degrees of freedom than the corresponding 1995 method 1. The 1996 variance estimation design consists of collapsed strata that may introduce a much larger stratum-collapse bias than occurred in 1995, and furthermore, the PSUs within each 1996 collapsed stratum have greater PSU weight diversity than in 1995 which may reduce stability.

The section on SUBSETTED DATA ANALYSES in the 1995 section should be read considering the changes provided in this 1996 section.