

## **NESARC DATA NOTES**

These Data Notes provide information regarding the NESARC Wave 1 Public Use Data File that may be helpful to Public Use Data File users. This section will be updated from time to time to incorporate further data notes of interest.

Revised: 7/20/2004

## General Data Notes

1. Variable names are derived from a combination of section number and individual question number within sections. For example, responses to question 2 in section 4A are coded in a variable named S4AQ2. In some sections (notably sections 2B, 3A, 3C, and 12, covering alcohol, tobacco, drug, and betting experiences, respectively), the questionnaire includes a number of columns for each item. In this type of layout the variable names include a letter designating the particular column. For example, data for the question asked in column B of item 21 under question 1A of section 2B are contained in a variable called S2BQ1B21. In section 3C, which asks questions about anywhere from 1 up to 10 specific medicine/drug types, the variable name also indicates to which of the 10 drug types the question applies. In this case, the variable called S3CD6Q14E11 provides data on the 6<sup>th</sup> drug category (cocaine) in column E of item 11 under question 14A of section 3C. (Please refer to a copy of the NESARC Wave 1 questionnaire to clarify variable naming conventions.)
2. Every effort was made to reduce respondent burden by skipping respondents past questions whose responses could be ascertained on the basis of answers to prior questions. For example, alcohol-related problems were first asked on a lifetime timeframe e.g., In your ENTIRE LIFE, did you ever find that your usual number of drinks had much less effect on you than it once did?. Respondents who answered affirmatively were then asked “Did this happen in the last 12 months?” and “Did this happen BEFORE 12 months ago?” If a respondent was classified as a former drinker (i.e., had not consumed any drinks in the past year), then his or her responses to the questions on whether alcohol-related problems ever happened were automatically copied into the corresponding fields for whether the problems happened before 12 months ago without actually asking the questions for that time period. **For the data user, this means that the definition of blank values for various survey items cannot be determined with certainty by looking at the skip instructions on the questionnaire. Rather, users should look at the codebook for definitions of blanks.**
3. Questions asking for durations (or periods of time over which something was happening) allowed respondents to answer in whatever time unit was most convenient. For example, a respondent might answer in weeks, months, or years. In the data file, all durations have been recoded and expressed in the smallest time unit associated with the particular question. Variable names for these recoded durations end in the letter “R”. The original variables associated with durations (e.g., number of weeks, number of months, etc.) are not included in the data file.
4. Sections 2B, 3A, 3C, 4A, 4C, 5, 6, 7, 8, 9, and 12 ask about substance use/mental health disorders which may happen once or more than once in a person’s life. The questionnaire is structured in such a way that age and duration questions asked of single-episode

respondents are different from those asked of multi-episode respondents. For example, multi-episode respondents are asked for their age at onset of first episode and age at onset of most recent episode. Single-episode respondents are asked only the age at first episode. For single-episode respondents the first episode is the same as the most recent episode. In the data file, the response to age at onset of first episode is placed in the multi-episode variables for age at onset of first or most recent episode. A similar recode is performed so that duration of “that” episode (for single-episode respondents) is coded into the multi-episode variables for duration of longest or only episode. Variable names for these recoded duration and age variables are based on the multi-episode variable names with an “R” appended. The original single-episode and multi-episode variables are not included in the data file.

5. The reliability and validity of the NESARC survey instrument, the Alcohol Use Disorder and Associated Disabilities Interview Schedule--DSM-IV Version (AUDADIS-IV) diagnostic and other measures can be found in the following published articles:
  - (1) Grant BF, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry, in press, 2004.
  - (2) Grant BF, Dawson DA, Hasin DS. The Alcohol use Disorder and Associated Disabilities Interview Schedule-DSM-IV Version. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2001.
  - (3) Canino GJ, Bravo M, Ramirez R, Febo V, Fernandez R, Hasin D. The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability and concordance with clinical diagnoses in a Hispanic population. J Stud Alcohol. 1999; 60:790-799.
  - (4) Chatterji S, Saunders JB, Vrsti R, Grant BF, Hasin DS, Mager D. The reliability of the Alcohol Use Disorders and Associated Disabilities Interview Schedule-Alcohol/Drug-Revised (AUDADIS-ADR) in India, Romania and Australia. Drug and Alcohol Depend. 1997; 47:171-185.
  - (5) Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug Alcohol Depend. 2003; 71: 7-16.
  - (6) Grant BF, Harford TC, Dawson DA, Chou PS, Pickering R. The Alcohol Use Disorder and Associated Disabilities Schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. Drug Alcohol Depend. 1995; 39:37-44.
  - (7) Hasin D, Carpenter KM, McCloud S, Smith M, Grant BF. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a clinical sample. Drug Alcohol Depend. 1997; 44:133-141.
  - (8) Vrsti, R., Grant BF, Chatterji S, Ustun BT, Mager D, Olteanu I, et al. The reliability of the Romanian version of the alcohol module of the WHO Alcohol Use Disorder and

Associated Disabilities Interview Schedule-Alcohol/Drug-Revised (AUDADIS-ADR). European Addiction Res. 1997; 40: 89-97.

- (9) Grant BF. DSM-III-R and ICD-10 alcohol and drug abuse/harmful use and dependence, United States, 1992: a nosological comparison. Alcohol Clin Exp Res. 1996; 21: 79-84.
- (10) Grant BF. DSM-III-R and proposed DSM-IV alcohol abuse and dependence, United States 1988: A nosological comparison. Alcohol Clin Exp Res. 1992; 16:1068-1075.
- (11) Grant BF. DSM-IV, DSM-III-R and ICD-10 alcohol and drug abuse/harmful use and dependence, United States, 1992: a nosological comparison. Alcohol Clin Exp Res. 1996; 20: 1481-1488.
- (12) Grant BF. The relationship between ethanol intake and DSM-III-R alcohol dependence: results of a national survey. J Subst Abuse. 1996; 5:257-267.
- (13) Grant BF, Harford TC. The relationship between ethanol intake and DSM-III-R alcohol dependence. J Stud Alcohol. 1990; 51: 448-456.
- (14) Grant BF, Harford TC. The relationship between ethanol intake and DSM-III alcohol use disorders: a cross-perspective analysis. J Subst Abuse. 1989; 1:231-252.
- (15) Hasin DS, Grant BF. Draft criteria for alcohol use disorders: comparison to DSM-III-R and implications. Alcohol Clin Exp Res. 1994; 18:1348-1353.
- (16) Hasin DS, Grant BF. Nosological comparisons of DSM-III-R and DSM-IV alcohol abuse and dependence in a clinical facility: comparison to National HIS88 results. Alcohol Clin Exp Res. 1994;18: 272-279.
- (17) Hasin D, Grant BF, Cottler L, Blaine J, Towle L, Ustun B, Sartorius N. Nosological comparisons of alcohol and drug diagnoses: a multisite, multi-instrument international study. Drug Alcohol Depend. 1997; 47:217-226.
- (18) Hasin D, Li Q, McCloud S, Endicott J. Agreement between DSM-III, DSM-III-R, DSM-IV and ICD-10 alcohol diagnoses in a US community-sample of heavy drinkers. Addiction. 1996; 91:1517-1527.
- (19) Hasin DS, Van Rossem R, McCloud S, Endicott J. Alcohol dependence and abuse diagnoses: validity in a community sample of heavy drinkers. Alcohol Clin Exp Res. 1997; 21:213-219.
- (20) Cottler LB, Grant BF, Blaine J, Mavreas V, Pull CB, Hasin D, Compton WM, Rubio-Stipee M, Mager D. Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. Drug Alcohol Depend. 1997; 47:195-205.
- (21) Hasin DS, Paykin A. Alcohol dependence and abuse diagnoses: concurrent validity in a nationally representative sample. Alcohol Clin Exp Res. 1999; 23:144-150.
- (22) Pull CB, Saunders JB, Mavreas V, Cottler LB, Grant BF, Hasin DS, et al. Concordance between ICD-10 alcohol and drug use disorder criteria and diagnoses as measured by the AUDADIS-ADR, CIDI and SCAN: results of a cross-national study. Drug Alcohol Depend. 1997; 47:207-216.
- (23) Hasin DS, Muthen B, Grant BF. The dimensionality of DSM-IV alcohol abuse and dependence: factor analysis in a clinical sample. Drug Alcohol Depend. 1993; 88:1079-1090.

- (24) Muthen B, Grant BF, Hasin DS. The dimensionality of alcohol abuse and dependence: factor analysis of DSM-III-R and proposed DSM-IV criteria in the 1988 National Health Interview Survey. Addiction. 1993; 88:1079-1090.
- (25) Nelson CB, Rehm J, Usten B, Grant BF, Chatterji S. Factor structure for DSM-IV substance disorder criteria endorsed by alcohol, cannabis, cocaine and opiate users: results from the World Health Organization Reliability and Validity Study. Addiction. 1999; 94:843-855.
- (26) Harford TC, Grant BF. Prevalence and population validity of DSM-III-R alcohol abuse and dependence: the 1989 National Longitudinal Survey on Youth. J Subst Abuse. 1994; 6:37-44.
- (27) Ustun B, Compton W, Mager D, Babor T, Baiyewu O, Chatterji S, et al. WHO study on the reliability and validity of the alcohol and drug use disorder instruments: overview of methods and results. Drug Alcohol Depend. 1997; 47: 161-170.
- (28) Hasin DS, Schuckit MA, Martin CS, Grant BF, Bucholz KK, Helzer JE. The validity of DSM-IV alcohol dependence: what do we know and what do we need to know. Alcohol Clin Exp Res. 2003; 27:244-252.
6. Informed Consent: All potential NESARC respondents were informed in writing about the nature of the survey, the statistical uses of the survey data, the voluntary aspect of their participation and the Federal laws that rigorously provide for the strict confidentiality of identifiable survey information. Those respondents consenting to participate after receiving this information were interviewed. The research protocol, including informed consent procedures, received full ethical review and approval from the U.S. Census Bureau and U.S. Office of Management and Budget.
7. The design effects characteristic of the NESARC require that you use special variance estimation statistical programs to derive standard errors. A number of such programs are available, including SUDAAN (which is the statistical software that we use here at NIAAA). For this software, we provide the following code necessary to specify the NESARC sample design:

```
PROC procname DESIGN=WR DATA=dsname;
NEST stratum psu / MISSUNIT;
WEIGHT weight;
```

The variables stratum, psu, and weight are described in the NESARC Wave 1 codebook.

## **Data Notes on Background Information (Section 1)**

1. Missing data were imputed for background variables deemed critical for analysis, e.g., age, sex, race, education, income, and so forth. Imputation was done by means of a hot deck procedure, except for age, sex, race and Hispanic origin, which also used logical checks. Hot decking was within categories defined by relevant characteristics. The “Source and Accuracy Statement for the 2001-2002 National Epidemiologic Survey on Alcohol and Related Conditions” contains a table listing the variables for which values were imputed, along with rates of imputation and the characteristics used for imputing each variable. For each imputed variable, the data file contains an associated “flag” variable that is coded with a value of “1” if the value of the variable was imputed. For example, for the variable S1Q1C (Hispanic or Latino origin) there is a flag variable called S1F1C. Flag variables enable the data user to identify cases where values were imputed and deal with the missing data in ways other than imputation, if so desired.
  
2. In 1997, the Office of Management and Budget issued revisions to its “Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting.” These revisions contained 2 major changes:
  - (1) Race was to be reported in one of five categories: (1) American Indian and Alaska Native; (2) Asian; (3) Black or African American; (4) Native Hawaiian and Other Pacific Islander; and (5) White. There would be no “other” race category.
  - (2) In surveys or other Federal data collection activities, persons asked about their race should be allowed to classify themselves as multi-racial by choosing as many of the five race categories as seem appropriate.

For analytical purposes, the Census Bureau has developed an algorithm to code a single race category for those individuals who identify themselves as multi-racial. When more than one race classification applies to the same individual a single race is selected from all chosen in the following order of preference:

- (1) Black or African American
- (2) American Indian and Alaska Native
- (3) Native Hawaiian and Other Pacific Islander
- (4) Asian
- (5) White

Thus, an individual who chooses Black and Asian will be classified as Black. An individual choosing American Indian, Native Hawaiian and White will be classified as American Indian.

The data file contains a race-ethnicity variable called ETHRACE2A (listed at the end of the codebook) which was constructed from the single classification race recode and the Hispanic origin variable (S1Q1C) and has the following 5 values:

- (1) White, non-Hispanic
- (2) Black, non-Hispanic
- (3) American Indian/Alaska Native, non-Hispanic
- (4) Asian/Native Hawaiian/Other Pacific Islander, non-Hispanic
- (5) Hispanic, any race

3. The 12 items required to create the SF12-V2 physical and mental functioning summary scales were collected in Section 1. For respondents who reported a valid (non-missing) response for the item on general health (S1Q16) but missing values for one to four of the other component variables, the missing data were imputed on the basis of the response to S1Q16. For respondents who had missing data for S1Q16 but valid responses to all of the other component items, S1Q16 was imputed on the basis of their score for the remaining items. Imputation flag variables indicate cases with imputed values. Other missing values in excess of the limits just described were left unimputed and resulted in a missing value for any scales to which they would have contributed: The scales and items on which they are based are listed below:

Norm-based Physical Disability Scale (NBPCS):	S1Q16 – S1Q22
Norm-Based Mental Disability Scale (NBMCS):	S1Q16 – S1Q22
Norm-Based Physical Functioning Scale:	S1Q171 – S1Q172
Norm-Based Role Physical Scale:	S1Q181 – S1Q182
Norm-Based Bodily Pain Scale:	S1Q20
Norm-Based General Health Scale:	S1Q16
Norm-Based Vitality Scale:	S1Q212
Norm-Based Social Functioning Scale:	S1Q22
Norm-Based Role Emotional Scale:	S1Q191-S1Q192
Norm-Based Mental Health Scale:	S1Q211, S1Q213

All of the scales were derived using techniques described in Ware JE, Kosinski M, Turner Bowker DM, Gandek B. *How to Score Version 2 of the SF-12 Health Survey*, Lincoln RI: Quality Metrics, 2002. This results in norm-based scores with a standardized range (0 to 100) and mean (50), facilitating comparisons across populations.

## **Data Notes on Alcohol Consumption Variables (Section 2A)**

### *1. Drinking Status*

Drinking status, described in a variable called CONSUMER, classified respondents as current (past-year) drinkers, former drinkers, or lifetime abstainers on the basis of three screening questions. These categories are defined as follows:

Current drinker: Drank at least 1 alcoholic drink in the last 12 months;

Former drinker: Did not drink (or unknown if drank) at least 1 drink in last 12 months but did drink at least 1 drink in lifetime

Lifetime abstainer: Did not drink (or unknown if drank) at least 1 drink in lifetime

### *2. Beverage-Specific Questions*

The NESARC contained separate series of comparable questions for past-year consumption of four different types of alcoholic beverages: coolers, beer, wine and liquor (i.e., distilled spirits). Each series of questions contained items on:

- 1) overall frequency of drinking the beverage type,
- 2) typical size of drink in ounces,
- 3) usual quantity consumed,
- 4) largest quantity consumed,
- 5) frequency of consuming the largest quantity,
- 6) frequency of consuming 5+ drinks of the beverage,
- 7) beverage subtype (e.g., regular beer, malt liquor, light or reduced calorie beer, ice beer), and
- 8) location where the beverage was usually consumed.

### *3. Quantity and Frequency*

The questions on frequency of drinking used categorical response options, shown to the respondents on flashcards. Respondent were asked to report exact quantities in an open-ended format. The questions on quantity and frequency were cleaned to ensure internal consistency, e.g. that the largest quantity did not exceed the usual quantity, that the overall frequency was not less than the frequency of consuming the largest quantity or 5+ drinks, and so forth.

#### 4. Drink Size

To aid the respondents in estimating their usual size of drink, they were shown flashcards containing categorical response options, accompanied by life-sized photographs of common glasses, with and without ice, with lines for various fill levels that indicated the corresponding number of ounces. The sizes in ounces corresponding to various response options (see Appendix A) are as follows:

<u>Oz.</u>	<u>Category</u>	<u>Oz.</u>	<u>Category</u>
1.0	1 ounce or shot, shot of unspecified size	16.9	½ liter bottle
1.5	1½ ounces or shots	33.8	1 liter bottle
2.0	2 ounces or shots; double	50.7	1.5 liter bottle; magnum
3.0	3 ounces or shots, triple; 3-ounce glass	59.2	1.75 liter bottle
4.0	4 ounces or shots, 4-ounce glass	101.4	3 liter bottle; double magnum
5.0	5-ounce glass, can or bottle	185.9	5 to 6 liter bottle or box
6.0	6-ounce glass, can or bottle		
7.0	7-ounce glass, can or bottle	8.0	½ pint
8.0	8-ounce glass, can or bottle*	16.0	pint
9.0	9-ounce glass, can or bottle	25.6	fifth
10.0	10-ounce glass, can or bottle	32.0	quart
12.0	12-ounce glass, can or bottle	64.0	½ gallon
15.0	15-ounce glass, can or bottle	128.0	gallon
16.0	16-ounce glass, can or bottle		
18.0	18-ounce glass, can or bottle	16.0	mug
20.0	20-ounce glass, can or bottle; schooner	60.0	pitcher
23.0	22- to 25-ounce can or bottle	67.6	growler
32.0	32-ounce can or bottle		
40.0	40- to 45-ounce bottle	48.0	six-pack of pony-size beer bottles
64.0	64-ounce bottle	72.0	six pack of regular beer bottles
		96.0	six-pack of large beer bottles/cans
1.5	1 jigger		
3.0	2 jiggers	blank	Other
4.5	3 jiggers		
6.0	4 jiggers	blank	Unknown
1.7	50-milliliter mini bottle		
6.3	187- milliliter bottle		
12.7	375-milliliter bottle; half bottle/carafe, split		
25.4	750-milliliter bottle; regular size wine bottle; full carafe		

\* Includes the response “7 or 8 ounce (pony size) can, bottle or glass”

## 5. *Ethanol Content*

Ethanol content was derived from questions on subtype of beverage usually consumed and additional information on main brand that is not included on the data tape. If a brand name was provided, its actual ethanol content was used, and the value for beverage subtype was edited for consistency with brand. If no brand was provided, then the ethanol contents were set to the following average values for the beverage subtypes:

### Coolers:

Wine coolers	0.050
Malt-based coolers	0.050
Liquor-based coolers	0.050
Prepackaged cocktails	0.120
Unknown	0.050

### Beer:

Regular beer	0.050
Malt liquor	0.065
Lite or reduced calorie beer	0.042
Ice beer	0.055
Unknown	0.050

### Wine:

Regular wine	0.125
Champagne or sparkling wine	0.120
Fortified wine (including sherry, port, sake)	0.180
Low-alcohol fruit-flavored wine	0.070
Unknown	0.120

### Liquor:

80-proof liquor including brandy	0.400
100-proof liquor	0.500
Liqueurs or cordials	0.271
Unknown	0.400

## 5. *Questions on Consumption of All Types of Beverages*

In addition to the beverage-specific series of questions, two additional sets of questions asked about consumption of *all types* of alcoholic beverages: 1) during the past 12 months and 2) during the period of heaviest drinking. These series included all the same items as the beverage-specific series, with the exception of drink size and beverage subtype. For the questions on consumption of all beverage types in the past 12 months, an additional question on frequency of drinking 4+ drinks was asked of women only.

Respondents who reported that they had never consumed more alcohol than in the past 12 months were not asked the questions on period of heaviest drinking. Rather, their responses for the last 12 months for all beverages combined were transferred into the corresponding fields for period of heaviest drinking.

#### 6. *Average Daily Volume of Ethanol Intake*

At the end of the data tape, there is a measure of average daily ethanol intake for the last 12 months, ETOTLCA2, that was derived by summing beverage-specific volumes across the four individual beverage types, as follows:

First, all of the reported frequencies of drinking were converted to number of drinking days per year, using the midpoints of the categorical response options, e.g., 3-4 times a week =  $3.5 \times 52 = 182$ . (For respondents who did not drink the type of beverage in question, the frequency was set to zero.)

For respondents whose largest quantity of drinks was five or fewer, average daily volume of ethanol intake had two components:

- 1) the usual quantity times the frequency of drinking that quantity:  $Q_U \times F_U$ , where  $F_U$  = the overall frequency of drinking minus the frequency of drinking the largest quantity, and
- 2) the largest quantity times the frequency of drinking the largest quantity:  $Q_L \times F_L$ .

The sum of these two products, representing the total number of drinks consumed per year, was then multiplied by the ethanol content of the drink in ounces, derived by multiplying the size of drink times the ethanol content by volume. The resulting annual volume of ethanol intake was divided by 365 to yield average daily ethanol intake of the beverage in question. These volumes were then summed across beverages to yield the overall average daily volume of ethanol intake.

For respondents those whose largest quantity of drinks was six or more, average daily volume had three components:

- 1) the usual quantity times the frequency of drinking that quantity:  $Q_U \times F_U$ , where  $F_U$  = the overall frequency minus the frequency of drinking 5+ drinks,
- 2) an intermediate component,  $Q_5 \times F_5$ , where  $F_5$  = the frequency of drinking 5+ drinks minus the frequency of drinking the largest quantity and  $Q_5 = \exp((\log(\max(5, Q_U)) + \log(Q_L - 1)) / 2)$ , i.e., the geometric mean of the band of quantities between 5 and the largest number of drinks, and
- 3) the largest quantity times the frequency of drinking the largest quantity:  $Q_L \times F_L$ .

Again, this sum of products was multiplied by the ethanol content per drink (see above) and divided by 365 to yield average daily ethanol intake of the beverage in question, and volumes were summed across beverages to yield the overall average daily volume of ethanol intake.

The approach described above was used for the great majority of cases. Rare exceptions were as follows:

If volume could not be calculated for one or more beverage types, then the average daily volume of ethanol intake was instead based on the maximum of either the sum of the known beverage-specific volumes or the volume calculated from the series of questions on consumption of all types of alcoholic beverages.

If the usual and largest quantities differed but their associated frequencies were the same, then the upper value of the frequency range was used for overall frequency of drinking and the lower value of the frequency range was used for the frequency of drinking the largest quantity. (This was only permitted for two frequency ranges, 3 to 6 times in the last year and 1 or 2 times in the last year.)

If the usual and largest quantities were the same, but the frequencies differed (e.g., people who usually drank 2 drinks, but sometimes drank 1 drink), then the quantity used in the first component of the volume estimation was set equal to the reported usual quantity minus 1.

**Tobacco and Medicine Use Data Notes  
(Sections 3A and 3B)**

1. The tobacco use section asks questions separately about the following 5 different types of tobacco: (1) cigarettes; (2) cigars; (3) pipe; (4) snuff; and (5) chewing tobacco. There is a variable called SMOKER that summarizes the overall tobacco use status of each person in the NESARC. The three values for SMOKER are:
  - (1) Current user (past 12 months) of one or more types of tobacco
  - (2) Ex-user (not in past 12 months) of one or more types of tobacco
  - (3) Lifetime non-user of any type of tobacco
  
2. The medicine use section asks questions separately about the following 9 different types of medicine or drug: (1) sedatives; (2) tranquilizers; (3) painkillers (opiates not including heroin or methadone); (4) stimulants; (5) marijuana; (6) cocaine or crack; (7) hallucinogens; (8) inhalants or solvents; and (9) heroin. There is also a tenth “other drug” category. There is a variable called DGSTATUS that summarizes the overall drug use status for each person in the NESARC. The three values for DGSTATUS are:
  - (1) Current user (past 12 months) of any type of drug
  - (2) Ex-user (not in past 12 months/unknown if in past 12 months) of any type of drug
  - (3) Lifetime non-user of any type of drug
  
3. A second variable in the medicine/drug use section summarizes overall drug use in a different way. CKSTATUS indicates the time frame for use of any type of drug and has the following five possible values:

1	Used in last 12 months only
2	Used prior to last 12 months only
3	Used in both time periods
9	Unknown when used
blank	Did not use in lifetime

**Mental Health Data Notes**  
**(Sections 4A, 4C, 5, 6, 7, 8, 9)**

1. The first 1, 2, or 3 questions in a section serve as screener questions that are asked of every respondent and determine which respondents go on to subsequent questions on individual symptom items in each section. (There are no screener questions for section 8, specific phobia.)
  
2. In some sections listing disorder symptoms there are numbered boxes associated with groups of individual symptom items. These boxes link individual symptoms to DSM-IV criteria. For example the first 4 symptom items under section 4A (major depression) are different indicators of 1 of 7 DSM-IV criteria for major depression, specifically, changes in weight or appetite. These boxes appear in some symptom lists as holdovers from a time before the AUDADIS was computerized and interviewers worked with a paper and pencil instrument. Some check items ask for box counts to control skip patterns during the course of the interview. Skip patterns are now controlled by computer during the course of an interview and counting of boxes goes on in the background, completely invisible to the interviewer. Check items on the paper questionnaire of some sections (e.g., section 12 on betting) ask for box counts when no boxes are indicated on the paper instrument. **This is an example of why the paper instrument provided to data users should be used only to view the wording of various questions and general structure of the questionnaire and should not be used to conduct an interview. In the computerized version of the AUDADIS there are skip patterns and response-contingent rewording of questions that are not indicated on the paper questionnaire.**

**Diagnostic Classifications**  
**(Sections 2b, 3a, 3b, 4a, 4c, 5, 6, 7, 8, 9, 10, 11a, 12)**

1. All diagnoses in the Wave 1 NESARC are made according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV: American Psychiatric Association). It would be helpful to obtain a copy of the DSM-IV, as it should help data users to understand how the diagnoses were constructed.
2. The substance use disorder diagnoses assessed in the Wave 1 NESARC include DSM-IV alcohol abuse and dependence and drug-specific diagnoses of abuse and dependence for nine categories of drugs (i.e., sedatives tranquilizers, opiates (other than heroin or methadone), stimulants, hallucinogens, cannabis, cocaine (including crack cocaine), inhalants/solvents, and heroin), along with an additional “other drug” category.
3. The DSM-IV mood and anxiety disorders assessed in the Wave 1 NESARC include major depression, dysthymia, manic episode, hypomanic episode, panic disorder with agoraphobia, panic disorder without agoraphobia, agoraphobia without history of panic disorder, social phobia, specific phobia, and generalized anxiety disorder.
4. The DSM-IV personality disorders assessed in the Wave 1 NESARC include avoidant personality disorder (PD), dependent PD, obsessive-compulsive PD, paranoid PD, schizoid PD, histrionic PD, and antisocial PD. All PD diagnoses are lifetime diagnoses.
5. DSM-IV pathological gambling and tobacco dependence were also assessed.
6. All diagnoses are provided on the public use data file for two time frames: (1) past 12 months; and (2) prior to the past 12 months. [The diagnostic variables appear at the end of the codebook and are well annotated.] We have provided a measure for lifetime diagnoses for selected disorders.
7. For mood and anxiety disorders in each of the above-reference time periods, four different diagnoses appear on the data file: (1) non-hierarchical diagnoses (i.e., those that do not use the exclusionary criteria of the DSM-IV); (2) those that exclude specific mood or anxiety disorders that are substance-induced; (3) those that exclude specific mood or anxiety disorders that are due to a general medical condition; and (4) those that exclude specific mood or anxiety disorders that are either substance-induced or due to a general medical condition.

In general, we use the diagnoses for mood and anxiety disorders that rule out substance-induced episodes or episodes due to a general medical condition. (Depending on your own analytic goals, you may wish to use the other diagnoses.) We refer to these diagnoses as “independent” and we differentiate them from substance-induced (including those due to general medical conditions) as follows:

Independent and substance-induced disorders were defined for respondents who met

criteria for specific mood and anxiety disorders for both the last 12 month and prior to the last 12 month timeframes. Disorders were classified as independent if: (1) the respondent abstained from alcohol and drug use in the time period; or (2) the episode(s) did not occur in the context of alcohol or drug intoxication or withdrawal; or (3) the episode(s) occurred prior to alcohol or drug intoxication or withdrawal; or (4) the episode(s) began after alcohol or drug intoxication or withdrawal, but persisted for more than one month after the cessation of acute alcohol or drug intoxication or withdrawal. Substance-induced disorders were defined as episodes that began after alcohol and/or drug intoxication and/or withdrawal, but were either: (1) not associated with a period of at least one month of abstinence; or (2) did not persist for more than one month after the cessation of acute alcohol or drug intoxication or withdrawal.

Respondents were classified with an independent mood or anxiety disorder in each time frame if none or only some of their episodes were substance-induced. Respondents were classified with a substance-induced disorder if all of their episodes in each time frame were substance-induced.

The prevalence of major depression reported here ruled out bereavement. All mood and anxiety disorders due to general medical conditions also were ruled out. The latter were defined as those occurring during the time frame when the respondent was physically ill or getting over being physically ill, with the additional requirement that a doctor or other health professional confirmed that the episode was related to the respondent's physical illness or medical condition. This definition also required the onset of the mood or anxiety disorder to begin during the time of a physical illness or during recovery from it.

8. We derive two diagnoses from Section 5 (High Mood). These are manic episode and hypomanic episode. These diagnostic groups are mutually exclusive.
9. We have derived diagnoses for: (1) panic disorder with agoraphobia; (2) panic disorder without agoraphobia; and (3) agoraphobia without a history of panic disorder. These three diagnoses are mutually exclusive.
10. With regard to all specific psychiatric disorders assessed in the NESARC, it is important to note that there will always be more individuals with data on onset, recency, duration, and number of episodes than there are individuals with the disorder. This is because we collected onset and the other aforementioned information on all individuals who passed the symptom item questions, even though they didn't satisfy all the diagnostic criteria for a positive diagnosis. We did this so that milder, subthreshold disorders could be examined in analyses of the NESARC data.
11. The diagnoses for substance use disorders (i.e., alcohol and drug abuse and dependence) and mood, anxiety, and personality disorders have been constructed by NIAAA psychiatric epidemiologists and their colleagues. However, we have left all the basic data on the data file so that alternative diagnoses could be created by users of the NESARC data file.

12. With regard to tobacco dependence, the diagnoses provided on the data file are for dependence on any tobacco product (i.e., cigarettes, cigars, pipes, snuff and chewing tobacco). The basic data to construct dependence diagnoses specific to each tobacco product have been retained on the public use file.
13. With regard to the DSM-IV personality disorders (excluding antisocial personality disorder), individuals needed to report the requisite number of symptom criteria for each diagnosis and, in addition, at least one symptom must have been associated with social and/or occupational dysfunction.
14. Variables identified as Check Items on the questionnaire do not appear on the data file. These variables were used only during the cleaning of the data file.

**FIPS Code Appendix**  
**U.S. States**

01	Alabama	30	Montana
02	Alaska	31	Nebraska
04	Arizona	32	Nevada
05	Arkansas	33	New Hampshire
06	California	34	New Jersey
08	Colorado	35	New Mexico
09	Connecticut	36	New York
10	Delaware	37	North Carolina
11	District of Columbia	38	North Dakota
12	Florida	39	Ohio
13	Georgia	40	Oklahoma
15	Hawaii	41	Oregon
16	Idaho	42	Pennsylvania
17	Illinois	44	Rhode Island
18	Indiana	45	South Carolina
19	Iowa	46	South Dakota
20	Kansas	47	Tennessee
21	Kentucky	48	Texas
22	Louisiana	49	Utah
23	Maine	50	Vermont
24	Maryland	51	Virginia
25	Massachusetts	53	Washington
26	Michigan	54	West Virginia
27	Minnesota	55	Wisconsin
28	Mississippi	56	Wyoming
29	Missouri		

## Appendix A Size of Typical Drink

1	1 ounce or shot, shot of unspecified size	29	½ liter bottle
2	1½ ounces or shots	30	1 liter bottle
3	2 ounces or shots; double	31	1.5 liter bottle; magnum
4	3 ounces or shots, triple; 3-ounce glass	32	1.75 liter bottle
5	4 ounces or shots, 4-ounce glass	33	3 liter bottle; double magnum
6	5-ounce glass, can or bottle	34	5 to 6 liter bottle or box
7	6-ounce glass, can or bottle		
8	7-ounce glass, can or bottle	35	½ pint
9	8-ounce glass, can or bottle*	36	pint
10	9-ounce glass, can or bottle	37	fifth
11	10-ounce glass, can or bottle	38	quart
12	12-ounce glass, can or bottle	39	½ gallon
13	15-ounce glass, can or bottle	40	gallon
14	16-ounce glass, can or bottle		
15	18-ounce glass, can or bottle	41	mug
16	20-ounce glass, can or bottle; schooner	42	pitcher
17	22- to 25-ounce can or bottle	43	growler
18	32-ounce can or bottle		
19	40- to 45-ounce bottle	44	six-pack of pony-size beer bottles
20	64-ounce bottle	45	six pack of regular beer bottles
		46	six-pack of large beer bottles/cans
21	1 jigger		
22	2 jiggers	47	Other
23	3 jiggers		
24	4 jiggers	99	Unknown
25	50-milliliter mini bottle (type sold on airlines)		
26	187- milliliter bottle (small individual wine bottle usually sold in 4-packs)		
27	375-milliliter bottle; half bottle of wine; half carafe; split		
28	750-milliliter bottle; regular size wine bottle; full carafe		

\*Includes the response “7 or 8 ounce (pony size) can, bottle or glass”