

**National Epidemiologic Survey on  
Alcohol and Related Conditions -III  
(NESARC-III)**

**Data Notes**

## **NESARC-III DATA NOTES**

These Data Notes provide information regarding the NESARC-III 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.

## 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 N4AQ2. In some sections (notably Sections 2B, 3A and 3C, covering alcohol, tobacco/nicotine, and drug 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 1 of Section 2B are contained in a variable called N2BQ1B21. In Section 3A, which asks about five different types of tobacco/nicotine products, these types are distinguished by a suffix of 1-5 at the end of the variable name. For example, the variable N3AQ41 provides data on Section 3A, question 4 for cigarettes (tobacco/nicotine type 1). In Section 3C, which asks questions about 10 specific medicine/drug types, the drugs are distinguished by the addition of Dx preceding the question number, where x = the drug type in question. In this case, the variable called N3CD6Q1E11 provides data on the 6<sup>th</sup> drug category (club drugs) in column E of item 11 under question 1 of Section 3C. This approach is also adopted in Section 3D, Drug Treatment. (Please refer to a copy of the NESARC-III 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. Similarly, if the respondent had started drinking in the past year (yes to N2AQ12B), 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 in the last 12 months. **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, 4B, 5, 6, 6A, 7, 8, 9, 10, 12, 17 and 18) 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 multiple-episode respondents. For example, multiple-episode respondents are asked for their age at onset of first episode and age at onset of most recent episode (age at recency), whereas single-episode respondents are asked only the age at first episode. In the data file, age at onset has been copied into the age at recency variable for single-episode respondents. Likewise, data on the duration of the only episode for single-episode respondents has been copied into the variable(s) for duration(s) of longest and/or most recent episodes for multiple-episode respondents. These merged variables may or may not have an R suffix following the multiple-episode variable name. The original single-episode variables are not included in the data file.
5. Several substance-related sections of the NESARC-III (2B, 3A, 3D) include questions on treatment that ask for the age when the respondent first received treatment and the age when they most recently did so. For the second of these questions, if the respondent indicated that he or she only received treatment once, then the age at first treatment was copied into the age at most recent treatment.
6. The reliability and validity of the diagnostic and other measures in NESARC-III survey instrument, the Alcohol Use Disorder and Associated Disability Interview Schedule--DSM-5 Version (AUDADIS-5), and its predecessors can be found in the following published articles:
  - (1) Goldstein RB, Chou SP, Smith SM, Jung J, Zhang H, Saha TD, Pickering RP, Ruan WJ, Huang B, Grant BF. Nosologic comparisons of DSM-IV and DSM-5 alcohol and drug use disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions-III, 2014, *manuscript submitted*.
  - (2) Grant BF, Amsbary M, Chu A, Sigman R, Kali J, Sugawana Y, Jiao R, Goldstein RB, Jung J, Zhang H, Chou PS, Saha TD, Huang B, Ruan WJ, Pickering RP, Smith SM. *Source and Accuracy Statement: National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III)*. National Institute on Alcohol Abuse and Alcoholism, Rockville, MD.
  - (3) Grant BF, Goldstein RB, Smith SM, Jung J, Zhang H, Chou SP, Pickering RP, Ruan WJ, Huang B, Saha TD, Aivadyan C, Greenstein E, Hasin DS. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5): reliability of substance use and psychiatric disorder modules in a general population sample, 2014, *Drug Alcohol Depend, in press*.

- (4) Hasin DS, Greenstein E, Aivadyan C, Stohl M, Aharonovich E, Saha TD, Goldstein RB, Grant BF. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-5 (AUDADIS-5): procedural validity of substance use disorders modules through clinical re-appraisal in a general population sample, 2014, *Drug Alcohol Depend*, in press.
- (5) Ruan WJ, Goldstein RB, Chou SP, Smith SM, Saha TD, Pickering RP, Dawson DA, Huang B, Stinson FS, Grant BF. The Alcohol Use Disorder and Associated Disabilities Interview Schedule – IV (AUDADIS-IV): reliability of new psychiatric diagnostic modules and risk factors in a general population sample. *Drug Alcohol Depend*. 2008;93:27-36.
- (6) 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*. 2004;61:361-368.
- (7) Grant BF, Dawson DA, Hasin DS. *The Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions Alcohol Use Disorder and Associated Disabilities Interview Schedule — DSM-IV Version*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2004.
- (8) 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.
- (9) 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.
- (10) 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.
- (11) 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.
- (12) Hasin DS, Paykin A. Alcohol dependence and abuse diagnoses: concurrent validity in a nationally representative sample. *Alcohol Clin Exp Res*. 1999; 23:144-150.
- (13) 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.
- (14) Chatterji S, Saunders JB, Vraiti 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.

- (15) 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.
- (16) 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.
- (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 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.
- (19) Vradi, 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.
- (20) 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.
- (21) 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.
- (22) 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.
- (23) 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.
- (24) 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.
- (25) 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.
- (26) 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.

- (27) 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.
- (28) 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.
- (29) 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.
- (30) 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.
- (31) 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.
- (32) 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.
- (33) Grant BF, Harford TC. The relationship between ethanol intake and DSM-III-R alcohol dependence. *J Stud Alcohol*. 1990; 51: 448-456.
- (34) 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.

7. Informed Consent: All potential NESARC-III 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 Westat Institutional Review Board and the Combined Neuroscience Institutional Review Board of the National Institutes of Health.

8. The design effects characteristic of the NESARC-III require that you use special variance estimation statistical programs that handle complex survey design and that generate the appropriate variance and standard error estimates. 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-III sample design:

```
PROC SORT DATA=dsname;
BY VARSTRAT VARUNIT; RUN;

PROC procname DESIGN=WR DATA=dsname;
```

```
NEST varstrat varunit / MISSUNIT;  
WEIGHT audweight;
```

In SAS 9.3 (requires version 12.1 of SAS/STAT), the comparable code is as follows:

```
PROC procname DATA=dsname VARMETHOD=TAYLOR;  
WEIGHT audweight;  
STRATA varstrat;  
CLUSTER varunit;
```

In STATA, the code is as follows:

```
SVYSET VARUNIT [PWEIGHT=audweight], STRATA(varstrat) VCE(LINEAR)
```

Note that the variables varstrat, varunit, and audweight are described in the NESARC-III codebook.

Also note that STATA is case sensitive with respect to variable names.

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

1. Missing data: The potential impact of residual item nonresponse in the NESARC, like most other surveys, is corrected through the use of imputation. Imputation rates are computed as the ratio of the number of eligible people who had a value imputed for that item to the number of all people eligible to respond to the item. The process by which values for missing or inconsistent data are determined is complex and varies by survey. In general, there is a continuum of certainty about the probable content of a missing data item. Analysts are confident about assigning values to a missing data item when related information is available on the same person record. For example, first name may be used to assign a value of sex. Such imputations are often known as assignments since they do not rely on data from a separate record. Confidence is lower when values for missing or inconsistent items cannot be derived from the same person record and must come from other respondents believed to have similar characteristics. When such donors are used, the item is said to have been allocated. Such imputation is considered less accurate than using information derived from the housing unit or person record. 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, N1Q16A (worked at a job or business in last 12 months) there is a flag variable called N1F16A. 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. In the NESARC-III, age and marital status variables were imputed using both assignment and allocation. All other variables imputed in the NESARC-III used only the allocation method. The imputation rates along with the characteristics constituting the donor cells for these variables are shown in the following table.

Table: Imputation of NESARC-III Variables

<b>VARIABLE</b>	<b>IMPUTATION RATE (%)</b>	<b>DONOR CELL CHARACTERISTICS</b>
Age (NAGE)	1.13	a. Assigned based on other reported age b. Allocation: sex and age interval
Marital status (NMARITAL)	0.06	Assignment method
Present situation ( N1Q12A1- N1Q12A9, N1Q12A10- N1Q12A14), imputed as a group	0.19	Worked in last 12 months, sex, age, Black race & Hispanic origin
Duration (weeks) of unemployment (N1Q12BR)	1.01	Ever worked, sex, age, Black race & Hispanic origin
Full-/Part-time student last year (N1Q13)	0.60	Age, employment status (FT, PT and other)
Educational attainment (NEDUC)	0.06	Sex, age, Black race & Hispanic origin
Worked in last 12 months (N1Q16A)	0.02	Age & sex
Ever worked in lifetime (N1Q16B)	0.02	Age & sex
Type of industry (N1Q17A)	0.52	Sex & educational attainment
Occupation (N1Q17B)	0.34	Sex, Black race & educational attainment
Type of Employer (N1Q17C)	0.66	Type of industry
Personal income (N1Q18A)	10.3	Age, sex, educational attainment, employment status (FT, PT and other) & occupation
Family income (N1Q19A)	11.5	Personal income, sex & adult relatives in household
Household income (N1Q20A)	13.1	Family income, age & adult non-relatives in household
Received social security in last 12 months (N1Q22A)	0.18	Age, ever worked & ever married
Received SSI in last 12 months (N1Q22B)	0.24	On disability (N1Q12A8), or limited activities due to physical health ( N1Q30A, N1Q30B, N1Q31A, N1Q31B)
Received TANF in last 12 months (N1Q22C)	0.17	Household income categories (N1Q20B), sex, Black race & children in household
Received WIC in last 12 months (N1Q22D)	0.13	Household income categories (N1Q20B), sex, pregnancy in last year, age & children in household
Received food stamps in last 12 months (N1Q23A)	1.18	Household income categories (N1Q20B), sex, age & children in household
Amount of food stamps received in last 12 months (N1Q23B)	5.29	Household income categories (N1Q20B), sex, age & children in household

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 or Alaska Native; (2) Asian; (3) Black or African American; (4) Native Hawaiian or 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 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 or Alaska Native
- (3) Native Hawaiian or 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 NETHRACE which was constructed from the single classification race recode (based on N1Q1F1 through N1Q1F5) and the Hispanic origin variable (N1Q1E), 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 (N1Q25) but missing values for one to four of the other component variables, the missing data were imputed on the basis of the response to N1Q25. For respondents who had missing data for N1Q25 but valid responses to all of the other component items, N1Q25 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 component items on which they are based are listed below:

Norm-based Physical Summary Scale (NNBPCS):	N1Q25, N1Q30A-N1Q35
Norm-Based Mental Summary Scale (NNBMCS):	N1Q25, N1Q30A-N1Q35
Norm-Based Physical Functioning Scale (NNBS1):	N1Q30A-N1Q30B
Norm-Based Role Physical Scale (NNBS2):	N1Q31A-N1Q31B
Norm-Based Bodily Pain Scale (NNBS3):	N1Q35
Norm-Based General Health Scale (NNBS4):	N1Q25
Norm-Based Vitality Scale (NNBS5):	N1Q33B
Norm-Based Social Functioning Scale (NNBS6):	N1Q34
Norm-Based Role Emotional Scale (NNBS7):	N1Q32A-N1Q32B
Norm-Based Mental Health Scale (NNBS8):	N1Q33A, N1Q33C

All of the scales were derived using the scoring methods 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-Metric, Incorporated, 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 NCONSUMER, 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 drank at least 1 drink in lifetime
- Lifetime abstainer: Did not drink (or unknown if drank) at least 1 drink in lifetime

### 2. *Questions for All Alcoholic Beverages Combined*

The NESARC-III contained an extensive series of questions for past-year consumption of any type of alcoholic beverage, i.e., all alcoholic beverages combined. This comprised items on:

- overall frequency of drinking,
- usual quantity consumed,
- largest quantity consumed, with a categorical probe if unknown/refused
- frequency of consuming the largest quantity,
- frequencies of consuming 4+ drinks and 4+ drinks within a 2-hour period, only asked of women and men 65 and older,
- frequencies of consuming 5+ drinks and 5+ drinks within a 2-hour period
- frequencies of consuming 8+ and 12+ drinks

An identical series of questions was asked for the self-reported period of heaviest drinking, for past-year drinkers who reported a time when they drank more heavily than in the past year and for all former drinkers. For past-year drinkers who did not report a period of drinking more than in the past year, the past-year values for these variables were copied into the corresponding fields for period of heaviest drinking. The questions for period of heaviest drinking included one additional item on the type of alcoholic beverage most often consumed during that period.

### 3. *Beverage-Specific Questions*

The NESARC-III 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:

- overall frequency of drinking the beverage type,
- typical size of drink in ounces,
- usual quantity consumed,
- largest quantity consumed,

- frequency of consuming the largest quantity,
- frequency of consuming 5+ drinks of the beverage,
- beverage subtype (e.g., regular beer, malt liquor, light or reduced calorie beer, ice beer),
- location where the beverage was usually consumed

#### 4. *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.

#### 5. *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:

1.0	1 ounce or shot, shot of unspecified size	1.7	50-milliliter mini bottle (type sold on airlines)
1.5	1½ ounces or shots	6.3	187- milliliter bottle (small individual wine bottle usually sold in 4-packs)
2.0	2 ounces or shots; double, 2-ounce can or bottle	12.7	375-milliliter bottle; half bottle of wine; half carafe; split
3.0	3 ounces or shots, triple; 3-ounce glass	25.4	750-milliliter bottle; regular size wine bottle; full carafe
4.0	4 ounces or shots, 4-ounce glass	16.9	1/2 liter bottle
5.0	5-ounce glass, can or bottle	33.8	1 liter bottle
6.0	6-ounce glass, can or bottle	50.7	1.5 liter bottle; magnum
7.0	7-ounce glass, can or bottle	39.2	1.75 liter bottle
8.0	8-ounce glass, can or bottle	101.4	3 liter bottle; double magnum
9.0	9-ounce glass, can or bottle	185.9	5 to 6 liter bottle or box
10.0	10-ounce glass, can or bottle	6.8	1/2 pint
12.0	12-ounce glass, can or bottle	16.0	Pint
15.0	15-ounce glass, can or bottle	25.6	Fifth
16.0	16-ounce glass, can or bottle	32.0	Quart
18.0	18-ounce glass, can or bottle	64.0	1/2 gallon
20.0	20-ounce glass, can or bottle; schooner	128.0	Gallon
23.0	22- to 25-ounce can or bottle	16.0	Mug
32.0	32-ounce can or bottle	60.0	Pitcher
40.0	40- to 45-ounce bottle	67.6	Growler
64.0	64-ounce bottle	48.0	Six-pack of pony-size beer bottles
1.5	1 jigger	72.0	Six pack of regular beer bottles
3.0	2 jiggers	96.0	Six-pack of large beer bottles/cans
4.5	3 jiggers		
6.0	4 jiggers		

#### 6. *Ethanol Content*

Ethanol content was derived from questions on subtype of beverage usually consumed and additional information on main brand that was 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 values for the beverage subtypes:

Coolers:

Wine/malt/liquor based coolers	0.050
Hard lemonade	0.050
Hard iced tea	0.060
Hard cider	0.060
Alcoholic energy drinks	0.080
Prepackaged cocktails	0.125
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	0.060
Unknown	0.125

Liquor:

80-proof liquor including brandy	0.400
>80 - 100-proof liquor	0.450
>100-proof liquor	0.750
Liqueurs or cordials	0.271
Unknown	0.400

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, NETOTLCA2, which 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 in ounces 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  $Q_5 = \exp((\log(\max(5, Q_U)) + \log q_{coolch-1})) / 2)$ , 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.

## 7. *Exceeding Drinking Guidelines*

Two past-year measures of exceeding the NIAAA Low-Risk Drinking Guidelines (add link to website) were derived from the NESARC-III data. The first was a dichotomous variable (NEXCEED) that was positive if the respondent exceeded either the recommended weekly limits (no more than 14 drinks a week for men and no more than 7 for women) or daily limits (no more than 4 drinks on any day for men and no more than 3 for women). The second, NFEXMAX, indicated the frequency of exceeding the daily limits.

In the derivation of NEXCEED, exceeding the weekly limits was based on average daily ethanol intake, i.e., whether NETOTACA2 exceeded 1.2 ounces for men (2 standard drinks of 0.6 ounces per day on average, corresponding to 14 drinks per week) or 0.6 ounces for women (1 standard drink of 0.6 ounces per day on average, corresponding to 7 drinks per week).

Exceeding the daily limits was based on two factors. First, we examined the usual and largest quantities of coolers, beer, wine and liquor, multiplied times their estimated ethanol content (e.g., times NCOOLETH, NBEERETH, NWINEETH or NLIQRETH). If any of these exceeded 2.7 ounces (4.5 standard drinks) for men or 2.1 ounces (3.5 standard drinks) for women, then the respondent was considered to have exceeded the daily limits. In addition, if men reported drinking 5+ drinks or women 4+ drinks with any non-zero frequency in the past year, they were considered to have exceeded the daily limits.

The second variable, NFEXMAX, reflected the largest of the frequencies associated with any of the conditions listed above for exceeding the daily limits. Thus if a man reported drinking 5+ drinks once a month but reported drinking four 12-ounce cans of malt liquor with an ethanol content of .065 ABV (for an ethanol intake of  $4 \times 12 \times .065 = 3.12$  ounces) once a week, then his frequency of exceeding the daily limits would be 52 times a year.

A second version of NEXCEED and NFEXMAX, called NEXCEED2 and NFEXMAX2, was derived as described above, but women's drinking limits were applied to men 65 years of age and older (as is recommended in the NIAAA Low-Risk Drinking Guidelines).

Past-year drinkers with missing values on any of the information needed to derive these variables were assigned missing values for the measures of exceeding the Drinking Guidelines. Former drinkers and lifetime abstainers were assigned negative values for the dichotomous variables and frequencies of zero.

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

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

**Mental Health Data Notes**  
**Mood, Anxiety, Post-Traumatic, and Eating Problems**  
**(Sections 4A, 4B, 5, 6, 6A, 7, 8, 9, 12, 17, 18)**

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-5 criteria. For example, items 3c-3f under section 4A (major depression) are different indicators of one of the nine DSM-5 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-5 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 traumatic experiences 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-5 there are skip patterns and response-contingent rewording of questions that are not indicated on the paper questionnaire.**

## **Diagnostic Classifications**

**(Sections 2b, 3a, 3c, 4a, 4b, 5, 6, 6a, 7, 8, 9, 10, 11a, 12, 17, 18)**

1. All diagnoses in the NESARC-III are made according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5: American Psychiatric Association). It would be helpful to obtain a copy of the DSM-5, as it should help data users to understand how the diagnoses were constructed.
2. The substance use disorder diagnoses assessed in the NESARC-III include DSM-5 alcohol use disorder, nicotine use disorder and drug-specific diagnoses of drug use disorder for nine categories of drugs (i.e., sedatives or tranquilizers, opiates (other than heroin), cannabis, cocaine (including crack cocaine), stimulants, hallucinogens, club drugs, inhalants/solvents, and heroin), along with an additional “other drug” category. DSM-IV alcohol and drug-specific abuse and dependence and DSM-IV nicotine dependence were also assessed.
3. The DSM-5 mood and anxiety disorders assessed in the Wave 1 NESARC include major depression, dysthymia, manic episode, hypomanic episode, panic disorder, agoraphobia, social phobia, specific phobia, and generalized anxiety disorder.
4. The DSM-5 personality disorders assessed in the NESARC-III include borderline personality disorder (PD), schizotypal PD, and antisocial PD. All PD diagnoses are lifetime diagnoses.
5. DSM-5 post-traumatic stress disorder and eating disorders (anorexia nervosa, bulimia nervosa, and binge-eating disorder) 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 measures of lifetime diagnoses for selected disorders.
7. For mood and anxiety disorders in each of the above-reference time periods, two different diagnoses appear on the data file: (1) non-hierarchical diagnoses (i.e., those that do not use the exclusionary criteria of the DSM-5); and (2) 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 in the last 12 month and/or 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 all 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.

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.
9. With regard to all specific psychiatric disorders assessed in the NESARC-III, 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-III data.
10. The diagnoses for substance use disorders, mood, anxiety, eating 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-III data file.
11. 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 or chewing tobacco, and e-cigarettes or e-liquid). The basic data to construct dependence diagnoses specific to each tobacco product have been retained on the public use file.
12. With regard to the DSM-5 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.

13. 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.

## **Appendix A**

### **Size of Typical Drink**

- 1 1 ounce or shot, shot of unspecified size
- 2 1½ ounces or shots
- 3 2 ounces or shots; double, 2-ounce can or bottle
- 4 3 ounces or shots, triple; 3-ounce glass
- 5 4 ounces or shots, 4-ounce glass
- 6 5-ounce glass, can or bottle
- 7 6-ounce glass, can or bottle
- 8 7-ounce glass, can or bottle
- 9 8-ounce glass, can or bottle
- 10 9-ounce glass, can or bottle
- 11 10-ounce glass, can or bottle
- 12 12-ounce glass, can or bottle
- 13 15-ounce glass, can or bottle
- 14 16-ounce glass, can or bottle
- 15 18-ounce glass, can or bottle
- 16 20-ounce glass, can or bottle; schooner
- 17 22- to 25-ounce can or bottle
- 18 32-ounce can or bottle
- 19 40- to 45-ounce bottle
- 20 64-ounce bottle
- 21 1 jigger
- 22 2 jiggers
- 23 3 jiggers
- 24 4 jiggers
- 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
- 29 ½ liter bottle
- 30 1 liter bottle
- 31 1.5 liter bottle; magnum
- 32 1.75 liter bottle
- 33 3 liter bottle; double magnum
- 34 5 to 6 liter bottle or box
- 35 ½ pint
- 36 Pint
- 37 Fifth
- 38 Quart
- 39 ½ gallon
- 40 Gallon
- 41 Mug
- 42 Pitcher
- 43 Growler
- 44 Six-pack of pony-size beer bottles
- 45 Six pack of regular beer bottles
- 46 Six-pack of large beer bottles/cans
  
- 47 Other