2012 NATIONAL SURVEY ON DRUG USE AND HEALTH

METHODOLOGICAL RESOURCE BOOK SECTION 16a: 2012 MENTAL HEALTH SURVEILLANCE STUDY DESIGN AND ESTIMATION REPORT

Substance Abuse and Mental Health Services Administration
Center for Behavioral Health Statistics and Quality
Rockville, Maryland 20857

June 2014

2012 MENTAL HEALTH SURVEILLANCE STUDY DESIGN AND ESTIMATION REPORT

Prepared for the 2012 Methodological Resource Book (Section 16a)

Contract No. HHSS283201000003C
RTI Project No. 0212800.001.112.008.006
Deliverable No. 39

RTI Authors:

Dan Liao
Phillip Kott
Jeremy Aldworth
Feng Yu
Rhonda Karg
Bonnie Shook-Sa
Teresa Davis

Project Director:

David Hunter

SAMHSA Project Officer:

Peter Tice

SAMHSA Authors:

Jonaki Bose
Joseph Gfroerer
Sarra Hedden

For questions about this report, please email Peter.Tice@samhsa.hhs.gov.

Prepared for Substance Abuse and Mental Health Services Administration, Rockville, Maryland

Prepared by RTI International, Research Triangle Park, North Carolina

June 2014

Recommended Citation: Center for Behavioral Health Statistics and Quality. (2014). *2012 National Survey on Drug Use and Health: Methodological Resource Book (Section 16a, 2012 Mental Health Surveillance Study: Design and Estimation Report)*. Substance Abuse and Mental Health Services Administration, Rockville, MD.

Acknowledgments

This report would not be possible without the guidance and input of staff from the Center for Behavioral Health Statistics and Quality (CBHSQ) within the Substance Abuse and Mental Health Services Administration (SAMHSA). In particular, CBHSQ's Art Hughes provided useful comments. Special thanks are also due to several staff members at RTI International (a trade name of Research Triangle Institute). James R. Chromy, Katherine B. Morton, Rachel Harter, and Kevin Wang provided valuable comments on the documentation of the sample design, sample weighting, and estimation methods. Dave Hunter is the RTI Project Director. Debbie Bond and Valerie Garner word processed and formatted the report, while Richard Straw copyedited it and Teresa Bass coordinated its Web production.

Table of Contents

Table of Contents (continued)

Table of Contents (continued)

List of Tables

List of Tables (continued)

List of Tables (continued)

List of Tables (continued)

List of Figures

List of Exhibits

1. Introduction

The overarching goal of the Mental Health Surveillance Study (MHSS) of the National Survey on Drug Use and Health (NSDUH) is to provide accurate estimates of the prevalence of serious mental illness (SMI) among adults aged 18 or older at the national and State levels. Public Law No. 102-321, the Alcohol, Drug Abuse, and Mental Health Administration Reorganization Act of 1992, established a block grant for U.S. States to fund community mental health services for adults with SMI. The law required States to include prevalence estimates in their annual applications for block grant funds. This legislation also required the Substance Abuse and Mental Health Services Administration (SAMHSA) to develop an operational definition of SMI and to produce national and State estimates. The MHSS clinical follow-up study was conducted to develop a model to generate estimates of SMI. However, the clinical data have the potential to be used for a variety of other purposes beyond this primary purpose, such as estimating other categories of mental illness (e.g., "mild," "moderate," or "any" mental illness).

On May 20, 1993, SAMHSA's Center for Mental Health Services (CMHS) published its definition of SMI in the *Federal Register*:

> Pursuant to Section 1912(c) of the Public Health Services Act, as amended by Public Law 102-321, "adults with serious mental illness" are defined as the following:
>
> - Persons aged 18 and over, who currently or at any time during the past year, have had diagnosable mental, behavioral, or emotional disorder of sufficient duration to meet diagnostic criteria specified within DSM-III-R [sic] that has resulted in functional impairment, which substantially interferes with or limits one or more major life activities.
> - These disorders include any mental disorders (including those of biological etiology) listed in DSM-III-R or their ICD-9-CM equivalent (and subsequent revisions), with the exception of DSM-III-R "V" codes, substance use disorders, and developmental disorders, which are excluded unless they co-occur with other diagnosable serious mental illness.
> - All of these disorders have episodic, recurrent, or persistent features; however, they vary in terms of severity or disabling effects. Functional impairment is defined as difficulties that substantially interfere with or limit role functioning in one or more major life activities, including basic daily living skills (e.g., eating, bathing, dressing); instrumental living skills (e.g., maintaining a household, managing money, getting around the community, taking prescribed medication); and functioning in social, family, and vocational/educational contexts.
> - Adults who would have met functional impairment criteria during the referenced year without benefit of treatment or other support services are considered to have serious mental illnesses.

In December 2006, a technical advisory group meeting of expert consultants was convened by the Office of Applied Studies (OAS, now the Center for Behavioral Health Statistics and Quality [CBHSQ]) and CMHS to solicit recommendations for mental health surveillance data collection strategies among the U.S. population. The panel recommended that

NSDUH be used to produce estimates of SMI among adults by including short scales in NSDUH's main interview that are strong predictors of SMI and that a "gold-standard" clinical psychiatric interview be administered on a subset of respondents to provide the data for estimating a statistical model that predicts SMI. In response, SAMHSA's CBHSQ initiated the MHSS as part of NSDUH to develop and implement a method to estimate SMI. At the time, NSDUH contained a six-item scale (Kessler-6 or K6) with five response options in each item that captured information on psychological distress in the past 12 months (Kessler et al., 2003). However, the K6 scale is not a diagnostic instrument and does not capture information on functional impairment, which is needed to determine whether a respondent can be categorized as having SMI under SAMHSA's definition. In consultation with the technical advisory group, two candidate impairment scales were selected by SAMHSA to be added to the 2008 NSDUH. They were an abbreviated version of the World Health Organization Disability Assessment Schedule (WHODAS; Rehm et al., 1999) and the Sheehan Disability Scale (SDS; Leon, Olfson, Portera, Farber, & Sheehan, 1997). An initial step was to modify these scales for use in a general population survey, including changes to question wording and length, which resulted in an abbreviated eight-item version of the WHODAS (Novak, Colpe, Barker, & Gfroerer, 2010). Further details of the K6 scale are given in Appendix A, and details of the two impairment scales are given in Appendices B and C.

The MHSS clinical interviews were conducted first in 2008. A split-sample design was used in the 2008 NSDUH, for which all adult respondents to the main NSDUH interview received the K6, but a random half of the sample received the WHODAS and the other half received the SDS. In addition, a subsample of approximately 1,500 adult NSDUH participants completed a follow-up clinical interview to provide data for developing models to estimate mental illness using the adult NSDUH data from the main interview. The randomization of the impairment scales was maintained within this clinical interview subsample, which is referred to in this report as the clinical sample, so that about half of the MHSS clinical sample participants were administered the WHODAS and half were administered the SDS (i.e., there were approximately 750 completed interviews from each half sample). Each participant in the 2008 MHSS clinical study was administered the Structured Clinical Interview for DSM-IV (SCI67D),[1] which was adapted for this study by mental health clinicians for paper-and-pencil interviewing administered over the telephone. The clinical interviews were administered approximately 2 to 4 weeks after the main NSDUH interview. Functional impairment ratings were assigned by clinical interviewers using the Global Assessment of Functioning (GAF) scale.[2] The modified SCID questionnaire for the 2008-2012 MHSS clinical study is shown in Appendix C of the 2008-2012 MHSS operations report (CBHSQ, 2014a). The model estimation analyses used gold-standard measures (i.e., the SCID/GAF combination as the indicator of SMI) in evaluating which

[1] The Structured Clinical Interview for the DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP) (First, Spitzer, Gibbon, & Williams, 2002).

[2] The GAF is a numeric scale (0 through 100) used to subjectively rate the social, occupational, and psychological functioning of adults and is presented and described in the DSM-IV-TR (see p. 32 of American Psychiatric Publishing, Inc., 2000; also see Endicott, Spitzer, Fleiss, & Cohen, 1976). Lower scores represent higher levels of functional impairment. Descriptions of impairment are provided at 10-point intervals (e.g., 1 to 10, 11 to 20, and so on up to 91 to 100). For example, a GAF score between 51 and 60 is described as having moderate symptoms of impairment, while a score higher than 60 represents several categories of impairment ranging from none to slight, and a score lower than 51 represents several categories ranging from serious to extreme.

combination of K6 and impairment scale worked best in the statistical model used to predict SMI status.

Based on an analysis of the 2008 MHSS data, it was determined that the WHODAS would be administered as the sole impairment scale in subsequent NSDUHs (starting in 2009) and that it would be used in combination with the K6 scale to predict SMI. For more details, refer to the 2008 MHSS analysis report by Aldworth et al. (2009). From 2009 through 2012, the MHSS was conducted similarly to the 2008 MHSS, except for two major differences: (1) only the WHODAS impairment scale was administered, and (2) the sample size was approximately 500 in 2009 and 2010, and the sample size was approximately 1,500 in 2011 and 2012.

The primary objective of the MHSS analysis is to produce annual national estimates of SMI prevalence that are accurate for all adults and for adult subpopulations. Secondary objectives include producing estimates of other categories of mental illness defined by level of impairment, such as mild (or low) mental illness (LMI), moderate mental illness (MMI), and any mental illness (AMI). These categories of mental illness, which are based on SCID disorder diagnoses and GAF scores, are defined in Table 1.1. A respondent was coded positive for SMI if he or she was determined to have any of the mental disorders (not including developmental or substance use disorders) assessed in the MHSS SCID *and* had a GAF score of 50 or below. AMI, defined as having a mental disorder regardless of the level of impairment due to that disorder, is the category obtained by collapsing the first three categories in Table 1.1 into a single category.

Table 1.1 Mental Illness Categories Defined by SCID Disorder Diagnosis and GAF Score: 2008-2012 MHSS Clinical Follow-Up Study

Mental Illness Category	SCID Disorder Diagnosis	GAF Score
Serious Mental Illness (SMI)	One or More	GAF $\leq$ 50 (severe or worse impairment)
Moderate Mental Illness (MMI)	One or More	50 < GAF $\leq$ 59 (moderate impairment)[1]
Mild (or Low) Mental Illness (LMI)	One or More	59 < GAF (at most mild impairment)
No Mental Illness (No MI)	None	GAF score not applicable

GAF = DSM-IV Axis V Global Assessment of Functional Scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition.

[1] DSM-IV description of moderate impairment based on GAF is 50 < GAF $\leq$ 60. The cutoff of 59 for MMI and LMI was chosen to conform to the corresponding cutoff selected by Kessler et al. (2003).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH MHSS clinical sample, 2008-2012.

A model was developed based on the 2008 clinical data and used in the 2008 through 2011 NSDUHs to produce a predicted probability of having SMI for each clinical interview respondent. A cut point was established among the fitted probabilities of having SMI based on the 2008 MHSS clinical data, such that adults with probabilities at or above the cut point were predicted to have SMI and the rest were not. Although this model was optimized to predict SMI, the SMI predicted probabilities were also used to predict MMI and LMI using different cut points (for more details about this model, see Liao et al., 2012).

With the accumulated MHSS clinical data collected from 2008 to 2012, SAMHSA determined that the 2008 model had some important shortcomings that had not been detected in the original model fitting because of the small number of respondents in the 2008 clinical sample. Specifically, the 2008 model substantially overestimated SMI and AMI among young adults in NSDUH relative to young adults in the clinical interview data. In addition, improvements were needed in the weighting procedures for the MHSS clinical data to account better for undercoverage and nonresponse (i.e., because only NSDUH respondents who answered their surveys in English were eligible for the clinical follow-up and because persons with mental illness appeared to be more likely to participate in the follow-up). Therefore, using the combined 2008-2012 clinical data, SAMHSA fit a more accurate model for the 2012 estimates with revised weights (subsequently referred to as the "2012 model" and "2012 estimation methods"). In particular, to reduce bias and improve prediction, additional mental health-related variables and an age variable were added in the 2012 model. In addition, to protect against potential coverage and nonresponse error, alternatives for the weights were applied to the clinical sample data for the model development. To provide consistent data for trend assessment, mental illness estimates for 2008 to 2011 using the new 2012 model were revised.

The remainder of this report is organized into seven chapters. Chapter 2 describes the mental illness and impairment scales and instruments that are used to produce SMI estimates. Chapter 3 discusses the sample design and methods for selecting respondents for the MHSS clinical interview. Chapter 4 illustrates how the collected MHSS data are processed for analytical purposes, including coding, editing, and imputation for the missing data. Chapter 5 clarifies the components of the MHSS analysis weights, including the methodology developed to adjust for nonresponse and noncoverage errors, while simultaneously preventing extreme weights. Chapter 6 summarizes the results of the descriptive analyses that compare the key demographic and psychosocial characteristics across samples from different time periods. Chapter 7 provides a general overview of the estimation methodology used to produce the prevalence estimates of mental illness among adults aged 18 or older and describes the revisions to the estimation methodology that were implemented in 2012.

2. Instruments Used in Measuring Mental Illness

2.1 Background

This chapter describes the psychological distress and impairment scales, as well as the clinical instrument, that were administered to adult respondents and used to produce estimates of serious mental illness (SMI). All adult respondents aged 18 or older in the National Survey on Drug Use and Health (NSDUH) were asked about their level of psychological distress. If a respondent had a psychological distress score greater than zero, he or she was then directed to questions on functional impairment.

In the main NSDUH interview, the Kessler-6 (K6) scale was used to assess psychological distress. In the 2008 NSDUH, a split-sample design was implemented in which all adult respondents in the main interview received the K6, but a random half of the sample received the World Health Organization Disability Assessment Schedule (WHODAS) and the other half received the Sheehan Disability Scale (SDS) for their impairment assessment. From the 2009 survey onward, the WHODAS was administrated as the sole impairment scale for adult respondents. The sample size was approximately 500 in 2009 and 2010 and approximately 1,500 in 2008, 2011, and 2012.

Each survey year from 2008 to 2012, a subsample of the adult NSDUH participants participated in a follow-up clinical interview and were administered the Structured Clinical Interview for DSM-IV (SCID) (First et al., 2002), including a module assessing Axis I disorders, and the Global Assessment of Functioning (GAF) scale by trained clinical interviewers.

2.2 Scales in the Main NSDUH Interview

2.2.1 Psychological Distress Scale (K6)

The K6 scale, used to capture nonspecific psychological distress (Kessler et al., 2003), consists of two sets of six questions that ask respondents how frequently they experienced symptoms of psychological distress during two different time periods: (1) during the past 30 days and (2) the one month in the past year when they were at their worst emotionally. Respondents were only asked about the second time period if they indicated that there was a month in the past 12 months when they felt more depressed, anxious, or emotionally stressed than they felt during the past 30 days. The six domains covered by the questions corresponded to how often the respondent felt (1) nervous, (2) hopeless, (3) restless or fidgety, (4) sad or depressed, (5) that everything was an effort, and (6) worthless. To create a score, the six items related to the first time period were coded from 0 to 4 so that "all of the time" was coded 4, "most of the time" 3, "some of the time" 2, "a little of the time" 1, and "none of the time" 0, with "don't know" and "refuse" also coded as 0. Summing across the six responses resulted in a total score with a range from 0 to 24. The six items related to the second time period were coded identically, and the worst K6 total score was calculated as the maximum of the total scores from the two time periods and is considered the past year K6 total score. An alternative version of the past year K6

total score was formulated as follows: past year K6 total scores of less than 8 were recoded as 0, and past year K6 total scores from 8 to 24 were recoded as 1 to 17. The reason for the alternative version was that the SMI prevalence was typically extremely low for respondents with past year K6 total scores of less than 8, and the prevalence rates were higher, in general, for total scores of 8 or greater. See Appendix A for the specific K6 scale items.

2.2.2 Functional Impairment Scales (WHODAS and SDS)

The abbreviated WHODAS, used to capture impairment data (Rehm et al., 1999), consists of eight questions that ask respondents how much their emotions, nerves, or mental health caused them to have difficulties in daily activities over the past year (Novak et al., 2010). Eight domains were covered by the following questions:

1. remembering to do things they needed to do,
2. concentrating on doing something important when other things were going on around them,
3. going out of the house and getting around on their own,
4. dealing with people they did not know well,
5. participating in social activities,
6. taking care of household responsibilities,
7. taking care of daily responsibilities at work or school, and
8. getting daily work done as quickly as needed.

To create a score, the eight items were coded from 0 to 3 so that "severe difficulty" was coded 3, "moderate difficulty" 2, "mild difficulty" 1, and "no difficulty" 0, with "don't know" and "refuse" also coded as 0. Some items had a fifth category to address "not applicable" responses. For example, the question about difficulties regarding taking care of daily responsibilities at work or school had a fifth category, "you didn't go to work or school." If this category was selected, then a further question was asked as to whether their emotions, nerves, or mental health caused them to be unable to go to work or school. A "yes" response to the follow-up question was coded 3, and a "no" response was coded 0.

One exception to this coding related to the last WHODAS item on how much difficulty respondents had in getting their daily work done as quickly as needed. This item was only asked if in the previous item on assessing their ability to function at work or school they responded to any of the first four categories (i.e., implying that they went to work or school) and was coded similarly to the other items. If they responded to the fifth category (i.e., that they did not go to work or school), their response to this item was determined by the final code for the follow-up item on whether their emotions, nerves, or mental health caused them to be unable to go to work or school.

Summing across the eight responses resulted in a total score with a range from 0 to 24. An alternative version of the WHODAS total score was formulated as follows: item scores of less than 2 were recoded as 0, and item scores from 2 to 3 were recoded as 1, then summed for a

total score ranging from 0 to 8. The reason for creating an alternative version of the WHODAS total score was the notion that using a dichotomous measure dividing respondents who experienced moderate or severe difficulties from the remaining respondents might produce a better predictor of SMI than would a linear continuous measure. See Appendix B for the actual questions used in the WHODAS.

The SDS, used only in half of the 2008 NSDUH sample to capture impairment data (Leon et al., 1997), consists of four questions that ask respondents how much their emotions, nerves, or mental health interfered with their daily activities over the past year. The following four domains were covered by the questions: (1) home management, (2) work, (3) close relationships with others, and (4) social life. For each of the four items, respondents were asked to select a number from 0 to 10 on a visual analog scale, where 0 means no interference, 1 to 3 mild interference, 4 to 6 moderate interference, 7 to 9 severe interference, and 10 very severe interference. Summing across the four responses resulted in a total score with a range from 0 to 40. An alternative version of the SDS total score was formulated as follows: item scores of less than 7 were recoded as 0, and item scores from 7 to 10 were recoded as 1, then summed for a total score ranging from 0 to 4. The alternative version of the SDS total score was again driven by the notion that a dichotomous measure dividing respondents who experienced severe or very severe interference from the remaining respondents might produce a better predictor of SMI than would a linear continuous measure. See Appendix C for the actual questions used in the SDS.

2.3 Clinical Follow-Up Instruments

Each participant in the 2008 to 2012 MHSS clinical follow-up study was administered standard clinical interview measures by mental health clinicians via paper-and-pencil interviewing over the telephone within 2 to 4 weeks of the NSDUH main interview. The MHSS clinical interview measure was the SCID (First et al., 2002), which is a semistructured diagnostic interview used to assess psychiatric disorders according to the criteria in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV; American Psychiatric Association, 1994). As a semistructured clinical interview, the SCID contains structured, standardized questions that are read verbatim and sequentially, combined with unstructured follow-up questions that the clinical interviewer tailors to the respondent based on clinical judgment and respondent reports. The SCID was modified for the MHSS to assess specific mental disorders and functioning in the past 12 months. The SCID was administered over the telephone by trained clinical interviewers. The SCID has been widely used in clinical calibration studies, such as the National Comorbidity Survey Replication (NCS-R; Kessler et al., 2004), the National Survey of American Life (Jackson, Neighbors, Nesse, Trierweller, & Torres, 2004), and NSDUH's substance use disorders reappraisal study (Jordan, Karg, Batts, Epstein, & Wiesen, 2008). It has demonstrated good reliability (Segal, Kabacoff, Hersen, Van Hasselt, & Ryan, 1995; Zanarini et al., 2000; Zanarini & Frankenburg, 2001) and validity (Fennig, Craig, Lavelle, Kovasznay, & Bromet, 1994; Kranzler, Kadden, Babor, Tennen, & Rounsaville, 1996; Kranzler et al., 1995; Ramirez Basco et al., 2000; Shear et al., 2000; Steiner, Tebes, Sledge, & Walker, 1995). Studies that compared telephone versus face-to-face administration of the SCID have found good agreement (Crippa et al., 2008; Hajebi et al., 2012; Kendler, Neale, Kessler, Heath, & Eaves, 1992; Kessler et al., 2004; Lee et al., 2008; Rohde, Lewinsohn, & Seely, 1997; Sobin et al., 1993).

Diagnostic modules in the MHSS version of the SCID are listed in Table 2.1. The assessment of lifetime manic episode was included to provide a context for understanding whether a past 12-month major depressive episode (MDE) was experienced as part of a unipolar mood disorder or as a component of a bipolar disorder (regardless of whether a manic episode also was experienced in the past year). The module to assess intermittent explosive disorder was obtained from the (optional) impulse control disorders section of the SCID. Although the module for substance use disorders was administered to respondents, substance use disorder is not included in SAMHSA's definition of SMI and was therefore not used in the estimation of SMI.

Table 2.1 Diagnostic Modules in the 2008-2012 MHSS SCID

MOOD DISORDERS	PAST YEAR EATING DISORDERS
Past Year Major Depressive Episode[1]	Anorexia Nervosa[1]
Lifetime Major Depressive Episode	Bulimia Nervosa[1]
Past Year Manic Episode[1]	
Lifetime Manic Episode	PAST YEAR IMPULSE CONTROL DISORDERS
Dysthymic Disorder[1]	Intermittent Explosive Disorder[1]
PAST YEAR PSYCHOTIC DISORDERS	
Psychotic Screen[1]	PAST YEAR SUBSTANCE USE DISORDERS
	Alcohol Abuse
PAST YEAR ANXIETY DISORDERS	Alcohol Dependence
Posttraumatic Stress Disorder[1]	Non-Alcohol Substance Abuse
Panic Disorder with and without Agoraphobia[1]	Non-Alcohol Substance Dependence
Agoraphobia without History of Panic Disorder[1]	
Social Phobia[1]	PAST YEAR ADJUSTMENT DISORDERS
Specific Phobia[1]	Adjustment Disorder[1]
Obsessive Compulsive Disorder[1]	
Generalized Anxiety Disorder[1]	

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition.

[1] Disorder used to determine gold-standard measures of serious mental illness and other categories of mental illness.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH MHSS clinical sample, 2008-2012.

In addition to the diagnostic modules, the MHSS SCID included four other sections:

1. An open-ended overview module is part of the standard SCID designed to elicit information about the respondent's diagnostic and treatment history and current status in a way that establishes some level of rapport between the clinical interviewer and the respondent.
2. The SCID screener instrument is a set of questions at the end of the overview section and is administered to all respondents. Its questions are taken from the body of the SCID and are the initial questions asked by the SCID for the disorders being assessed. These screening questions may help to reduce the potential effects of a "negative response bias" that may be especially problematic in the later sections of the SCID. Because of the structure of the SCID, there is a tendency for the respondent to notice that a "yes" answer to the initial probe question in a section results in follow-up

questions, whereas a "no" answer results in a skip to the next section, thus leading some subjects to give "no" answers as a way of speeding the interview along. By asking these questions up front and using the answers to these questions in the determination of whether a section should be skipped, response bias may be minimized.

3. With the DSM-IV Axis V GAF scale, the clinical interviewer was instructed to rate the respondent's period of worst psychological, social, and occupational functioning during the past year.
4. A section is included for documenting the clinical interviewer's impressions of the interview situation, including ratings of the respondent's level of privacy, cooperation, and comprehension, as well as the overall validity of the interview data (any interview deemed by the clinical interviewer or clinical supervision team to be of questionable validity was discarded).

For more details about these modules, see Colpe et al. (2010). Also, see Appendix C of the 2008-2012 MHSS operations report for the full modified SCID questionnaire (Center for Behavioral Health Statistics and Quality, 2014a).

3. Sample Design and Selection

3.1 Background

This chapter describes the sample design and methods for selecting respondents for the Mental Health Surveillance Study (MHSS) clinical interview. The clinical sample was selected from the main National Survey on Drug Use and Health (NSDUH) study sample of approximately 45,000 adults. The target population included those who completed the NSDUH main interview in English and excluded persons who completed the NSDUH main interview in Spanish. A probability sampling algorithm was programmed in the computer-assisted interviewing (CAI) instrument such that field interviewers (FIs) could, at the conclusion of the interview, recruit selected respondents for the subsequent clinical psychiatric interview conducted by telephone. This chapter describes, in detail, the annual MHSS clinical sample selection process, which began with a feasibility study in 2007 and evolved over the 2008 to 2012 data collection periods. Other issues related to the sample selection also are described. Descriptions of the 2008 to 2012 clinical samples are provided in Chapter 6.

3.2 Selecting the MHSS Clinical Follow-Up Samples

NSDUH respondents (i.e., NSDUH main interview respondents) aged 18 or older who completed their interviews in English were eligible to be sampled for the MHSS clinical follow-up study. The procedures used to identify which NSDUH respondents would be selected for the clinical interview varied over the data collection period. Prior to the initiation of the MHSS clinical follow-up study, a feasibility study was conducted in 2007 in which respondents were selected to complete the NSDUH main interview first, and all respondents who completed the main interview in English were selected for the clinical follow-up. Details about the 2007 feasibility study are discussed in Section 3.3.

In 2008 and 2009, the probability of selection for the clinical follow-up study was based on the respondent's Kessler-6 (K6) score. The K6 score was used because, as demonstrated by data from the National Comorbidity Survey Replication (NCS-R) clinical calibration study,[3] it is highly correlated with serious mental illness (SMI), the key characteristic of interest in the MHSS. Based on observed variation in the weights in the 2008 and 2009 MHSS clinical studies that affected the precision of resulting estimates, a design change was implemented starting in 2010. In 2010 through 2012, in order to decrease the variability of the weights, the probability of selection was based on the respondent's age, K6 score, and World Health Organization Disability Assessment Schedule (WHODAS) score. Because the target population and sampling frames for the 2008 through 2012 MHSS clinical follow-up studies remained the same, these changes in the allocation of selected persons did not affect the point estimates themselves. Instead, these design changes allowed for more precise estimates because they resulted in less variable analysis weights. Details of the selection probabilities for the MHSS clinical sample are discussed in Sections 3.4, 3.5, and 3.6.

[3] Kessler, R. C., attachment to a personal email communication to L. J. Colpe, August 1, 2007, Scidsmi-table-073107 (2) (2).doc.

3.3 2007 Feasibility Study

A feasibility study was conducted in June 2007 to test the procedures and instruments proposed for the 2008 MHSS clinical follow-up study. The feasibility study was conducted in order to assess all of the procedures necessary for conducting and reporting on the clinical interviews, as well as to assess the training needs for both the FIs and the clinical interviewers.

Unlike the 2008 through 2012 clinical studies, respondents to the clinical follow-up interview for the feasibility study were not selected from the main NSDUH study sample of adults. Instead, a new sample of dwelling units was selected from 10 sample segments of the 2006 NSDUH that had been retired from use in the main study. These retired segments were included in the 2005 and 2006 NSDUH samples and would not be reused in future NSDUH survey years. The segments were selected based on their proximity to experienced NSDUH FIs who were available to work on the feasibility study.

Consistent with the main NSDUH study, sampled dwelling units in the feasibility study were fielded and screened for main study eligibility. Within sampled dwelling units, zero, one, or two adults were selected for the initial NSDUH in-person interview. All respondents selected for and completing the initial interview in English were selected for the clinical follow-up telephone interview.

In preparation for the feasibility study, it was estimated that 270 dwelling units would be needed to yield 48 completed clinical interviews (see Table 3.1). Based on national-level experience from the 2006 NSDUH, adjusted for the shorter than normal data collection period,

Table 3.1 2007 Mental Health Surveillance Feasibility Study Design Parameters

Design Parameters	Count/Percent
Selected for the Initial Main Interview	
Selected Segments	10
Dwelling Units per Segment	27
Total Selected Dwelling Units	270
Dwelling Unit Eligibility Rate	83%
Eligible Dwelling Units	224
Expected Screening Response Rate	90%
Completed Screenings	202
Expected Yield 18 or Older	40%
Sampled Persons Aged 18 or Older	81
Assumed Interview Response Rate	70%
Selected for Telephone Clinical Follow-Up, by SPD Score	
Score 0 to 7	37, (66%)
Score 8 to 12	11, (19%)
Score 13 or Higher	8, (15%)
Total Selected for Telephone Clinical Follow-Up	56
Clinical Follow-Up Interview Response Rate	85%
Completed Clinical Follow-Up Interviews	48

NSDUH = National Survey on Drug Use and Health; SPD = serious psychological distress.

NOTE: Estimated sample proportions, eligibility rate, and response rates were derived from 2006 NSDUH data, adjusted for the shorter than normal data collection period.

an 83 percent dwelling unit eligibility rate, a 90 percent screening response rate, a 70 percent interview response rate, and an 85 percent clinical follow-up interview response rate were assumed. Also based on 2006 NSDUH data, it was expected that 66 percent of the interview respondents would have a serious psychological distress (SPD) score between 0 and 7, 19 percent would have a score between 8 and 12, and 15 percent would have a score of 13 or higher.

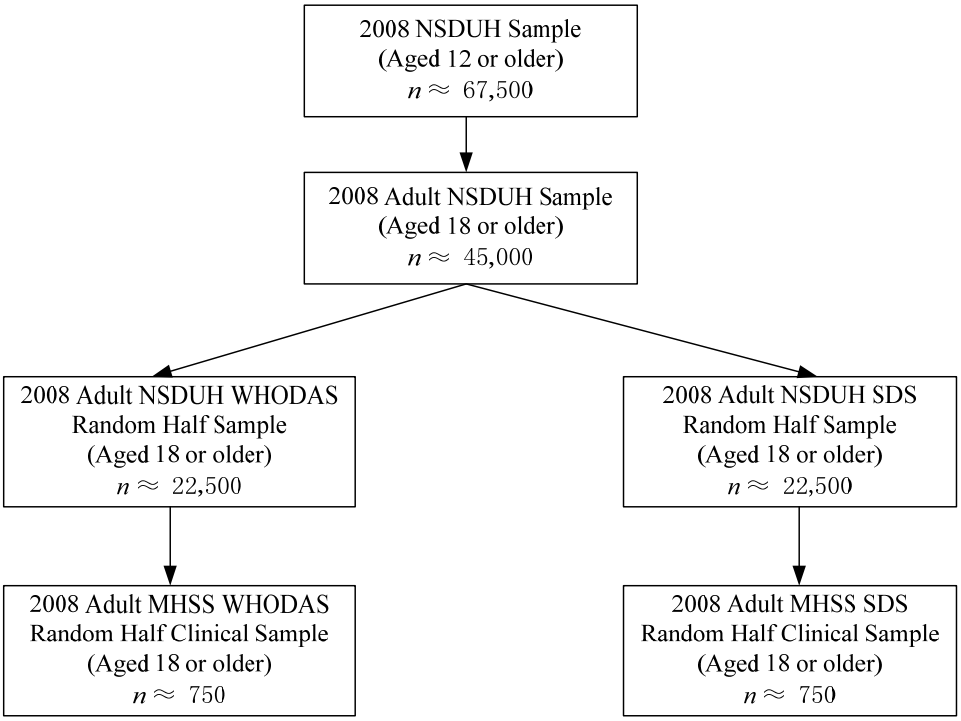
In the sample selection implementation, a total of 262 dwelling units were selected from the 10 sample segments. The dwelling unit sample size and person-level probabilities of selection were determined using sample allocation procedures identical to those for the 2007 NSDUH main study. State-specific eligibility rates, response rates, and yields were estimated using data from the 2006 NSDUH and adjusting for the shorter than normal data collection period. Using this State-specific information, fewer selected dwelling units were required than was originally anticipated. To ensure that the targeted number of clinical interviews was achieved with the feasibility study sample, a 20 percent reserve sample was also selected. To allow for the 20 percent reserve sample to be released as needed, the total sample of 262 dwelling units was partitioned into two releases within each segment.

A total of 232 sample dwelling units were fielded, 216 of which were eligible, giving a dwelling unit eligibility rate of 93.1 percent. Household screenings were completed at 195 households, for an unweighted screening response rate of 90.3 percent. A total of 102 sample persons were selected within these responding dwelling units, and interviews were completed with 70 for an unweighted initial interview response rate of 68.6 percent. Out of the 70 respondents who completed the initial main NSDUH field interview, 60 agreed to complete the follow-up clinical telephone interview. The unweighted percentage of respondents *agreeing* to complete the feasibility study clinical interview was 85.7 percent. Out of the 60 respondents who agreed to complete the clinical interview, 45 clinical interviews were completed. The unweighted percentage of respondents completing clinical interviews was 75.0 percent. The results of the feasibility study were used to inform the procedures utilized in the 2008 MHSS clinical study.

3.4 2008 Sample Allocation

The 2008 clinical data collection included 1,500 clinical follow-up interviews. In the 2008 NSDUH main study, adult respondents were randomly assigned to one of two functional impairment scales: an abbreviated version of the WHODAS (Novak et al., 2010; Rehm et al., 1999) and the Sheehan Disability Scale (SDS) (Leon et al., 1997). Approximately equal numbers of respondents received each of the functional impairment scales, and all of the respondents received the K6 items. The randomization of the impairment scales was maintained within the MHSS clinical sample so that about half of the MHSS clinical interview respondents (approximately 750) were administered the WHODAS and half were administered the SDS. A diagram illustrating the structure of the 2008 MHSS clinical study sampling design is given in Figure 3.1.

Figure 3.1 Structure of the 2008 Mental Health Surveillance Clinical Follow-Up Study Sampling Design



MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008.

NSDUH respondents eligible for the clinical follow-up were selected for the clinical interview through stratified random sampling. To achieve an efficient sample design, sampling strata were formed by grouping K6 scores with similar rates of SMI based on data collected from the NCS-R clinical calibration study. Several stratification options were considered in the optimization process, and a stratified sample design with seven sampling strata based on their K6 scores (0 to 3, 4 to 5, 6 to 7, 8 to 9, 10 to 11, 12 to 15, and 16 or higher) was adopted. This option was selected because it minimized the projected relative standard error (RSE) and more effectively smoothed out the sample distribution across K6 scores 4 through 11 compared with allocations based on fewer strata.

To optimize the MHSS clinical sample allocation for the clinical follow-up within scoring bands, assumed SMI rates were estimated using K6 scores and clinical case data from the NCS-R clinical calibration study.[4] Assumed SMI rates for the 2008 study were set equal to the NCS-R rates except in the instance where the K6 scores ranged from 0 to 7. In this instance, SMI rates were set lower under the assumption that fewer clinical positives would be identified in that scoring range. Population percentages by K6 group were estimated from the 2006 NSDUH.

[4] The design of the 2008 MHSS clinical study was based on data from the NCS-R clinical calibration study because these were the best estimates available at that time. Later designs based their assumptions on results from the prior year's MHSS clinical follow-up study.

Using Neyman's optimal allocation (Lohr, 1999), a solution that minimized the design effect for the prevalence of SMI in the clinical sample was selected. Table 3.2 shows the expected sample distribution for the 1,500 clinical follow-up interviews. The expected overall design effect[5] was 0.6363, which led to an expected effective sample size of 2,357. The projected standard error (SE) and RSE of the all-adult estimate of SMI prevalence based on the MHSS clinical sample were 0.59 and 6.57 percent, respectively.

Table 3.2 2008 Mental Health Surveillance Clinical Follow-Up Study Sample Allocation (*n* = 1,500)

K6 Score	Percent of Population	Assumed SMI Rate (%)	Sample Size
0 to 3	48.04	0.03	96
4 to 5	13.98	0.30	88
6 to 7	11.16	0.30	110
8 to 9	6.95	10.00	200
10 to 11	5.53	13.00	214
12 to 15	8.00	40.00	450
16 or Higher	6.34	67.00	343
Total	**100.00**	**8.95**	**1,501[1]**

K6 = Kessler-6, a 6-item psychological distress scale; NCS-R = National Comorbidity Survey Replication; NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness.

[1] This number does not equal the number for the sample size due to rounding.

NOTE: The population source is the 2006 NSDUH. Assumed SMI rates were estimated using data from the 2001-2002 NCS-R clinical calibration study.

The NSDUH CAI instrument included a sampling algorithm to indicate whether NSDUH main interview respondents had been sampled for the clinical follow-up study. FIs recruited sampled respondents for the subsequent clinical psychiatric interview that was conducted by telephone.

The probability sample of 1,500 clinical follow-up interviews was distributed across four calendar quarters with a slightly larger sample in the first quarter (425 follow-up interviews) and the remaining sample equally divided among the remaining quarters (approximately 358 interviews in each of the quarters 2 through 4 for a combined sample of 1,075 clinical follow-up interviews; see Table 3.3). The larger sample in quarter 1 was intended to provide some cushion should the clinical interview response rates be lower than anticipated. In addition, a slightly larger sample size in quarter 1 was needed to allow for preliminary analyses of the data. The sample sizes were determined based on an assumed 96 percent clinical interview eligibility rate,[6] an 85 percent agreement rate for the clinical follow-up interview, and a 90 percent participation rate among those who agreed to complete the interview.

[5] The expected design effect for the 2008 MHSS clinical follow-up study is the product of the usual design effect for adults in the main survey (about 3.0) and the design effect for the two-phase sample stratified by K6 scores (about 0.2).

[6] The MHSS clinical interview eligibility rate was not accounted for in the quarter 1 2008 design, but it was incorporated in the sample designs for quarters 2 to 4 in 2008.

Table 3.3 Design Parameters for the 2008 Mental Health Surveillance Clinical Follow-Up Study, Quarters 1 through 4

Design Parameter	Quarter 1 Total	Quarters 2 through 4 Total	Quarter 1 Expected Persons per Segment	Quarters 2 through 4 Expected Persons per Segment
Interview Respondents Aged 18 or Older, by K6 Score	11,250	33,750	6.3	6.3
Score 0 to 3 (42% of Cases)	4,725	14,175	2.6	2.6
Score 4 to 5 (13% of Cases)	1,463	4,388	0.8	0.8
Score 6 to 7 (12% of Cases)	1,350	4,050	0.8	0.8
Score 8 to 9 (8% of Cases)	900	2,700	0.5	0.5
Score 10 to 11 (6% of Cases)	675	2,025	0.4	0.4
Score 12 to 15 (10% of Cases)	1,125	3,375	0.6	0.6
Score 16 or Higher (9% of Cases)	1,013	3,038	0.6	0.6
Sampling Rate, by K6 Score				
Score 0 to 3	0.0075	0.0063		
Score 4 to 5	0.0223	0.0188		
Score 6 to 7	0.0301	0.0254		
Score 8 to 9	0.0823	0.0694		
Score 10 to 11	0.1174	0.0990		
Score 12 to 15	0.1481	0.1249		
Score 16 or Higher	0.1255	0.1058		
Selected for Telephone Clinical Follow-Up, by K6 Score				
Score 0 to 3	35	89	0.0	0.0
Score 4 to 5	33	82	0.0	0.0
Score 6 to 7	41	103	0.0	0.0
Score 8 to 9	74	187	0.0	0.0
Score 10 to 11	79	200	0.0	0.0
Score 12 to 15	167	422	0.1	0.1
Score 16 or Higher	127	321	0.1	0.1
Total Selected for Telephone Clinical Follow-Up	556	1,405	0.3	0.3
Percent Agreeing to Clinical Follow-Up	0.85	0.85		
Percent Completing the Clinical Follow-Up Interview	0.9	0.9		
Completed Clinical Follow-Up Interview, by K6 Score	425	1,075	0.2	0.2
Score 0 to 3	27	68	0.0	0.0
Score 4 to 5	25	63	0.0	0.0
Score 6 to 7	31	79	0.0	0.0
Score 8 to 9	57	143	0.0	0.0
Score 10 to 11	61	153	0.0	0.0
Score 12 to 15	127	322	0.1	0.1
Score 16 or Higher	97	246	0.1	0.0

K6 = Kessler-6, a 6-item psychological distress scale; NSDUH = National Survey on Drug Use and Health.

NOTE: The estimated distribution of interview respondents is based on the 2006 NSDUH.

Throughout the 2008 survey, the MHSS clinical sample was monitored, and the sampling parameters were modified on an as-needed basis to ensure that the goal of 1,500 completions was achieved. In addition, for the last 4 weeks in quarter 4,[7] the probability of selection of the NSDUH respondents for the clinical follow-up survey was set to zero so that cases would not be sampled without adequate time for completing the collection of the interviews (by December 19, 2008) (see Section 3.8.1 for further details).

[7] The MHSS clinical interview recruitment in 2008 ended on November 21st, the MHSS clinical study data collection ended on December 19th, and the NSDUH main study data collection ended on December 22nd.

At the end of the 2008 data collection period, 86 percent of the persons selected from NSDUH agreed to participate in the MHSS clinical follow-up, and 76 percent of those persons completed the MHSS clinical interview. The 2008 MHSS clinical study resulted in 1,500 completed clinical interviews, excluding cases removed from the dataset due to data errors (see Section 3.8.4) or unusable cases (see Section 3.8.5). A summary of the 2008 MHSS clinical interview respondents by quarter is included in Table 3.4.

Table 3.4 2008 Mental Health Surveillance Clinical Follow-Up Study, Quarters 1 through 4 Summary

Design Parameter	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
Interview Respondents Aged 18 or Older	10,692	12,816	11,355	10,815	45,678
Unweighted K6 Distribution, by K6 Score					
Score 0 to 3	0.45	0.45	0.45	0.44	0.45
Score 4 to 5	0.15	0.14	0.14	0.15	0.14
Score 6 to 7	0.10	0.10	0.10	0.10	0.10
Score 8 to 9	0.07	0.07	0.07	0.08	0.07
Score 10 to 11	0.05	0.06	0.06	0.06	0.06
Score 12 to 15	0.09	0.09	0.09	0.09	0.09
Score 16 or Higher	0.09	0.09	0.09	0.09	0.09
Eligible for MHSS	10,215	12,148	10,849	10,381	43,593
Eligibility Rate	*0.9554*	*0.9479*	*0.9554*	*0.9599*	*0.9544*
Selected for Telephone Clinical Follow-Up[1]	696	529	485	621	2,331
Zero Probability Cases	0	0	0	47	47
Agreed to Clinical Follow-Up	586	462	416	509	1,973
Percent Agreeing to Clinical Follow-Up (Including Zero Probability Cases)	*0.8420*	*0.8733*	*0.8577*	*0.8196*	*0.8464*
Percent Agreeing to Clinical Follow-Up (Excluding Zero Probability Cases)	*0.8420*	*0.8733*	*0.8577*	*0.8868*	*0.8638*
Completed Clinical Follow-Up Interviews	467	361	317	355	1,500
Clinical Follow-Up Interview Completion Rate	*0.7969*	*0.7814*	*0.7620*	*0.6974*	*0.7603*

K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

[1] Includes cases assigned a zero probability of selection that would have been selected based on their K6 rates.

NOTE: The overall response rate to the clinical follow-up study should also consider the response rate to the main NSDUH. In 2008, the weighted overall response rate for the main study was 65.79 percent.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008.

3.5 2009 Sample Allocation

The 2009 data collection was designed to yield 500 clinical follow-up interviews. A subsample of eligible respondents was selected for clinical follow-up with probabilities based on their K6 scores. Similar to the 2008 CAI instrument, the CAI instrument in 2009 included a sampling algorithm to indicate to an FI whether a NSDUH main interview respondent had also been selected for the clinical follow-up survey.

The 2009 sample was initially allocated to seven K6 scoring bands in the same proportions as the 2008 sample. Midway through the year, the decision was made to allocate the sample based on any mental illness (AMI) rather than SMI in order to reduce the selection

probabilities of respondents with extremely large weights while maintaining the efficiency of estimating SMI and AMI directly based on the clinical sample. A new allocation scheme was developed through Neyman's optimal allocation (Lohr, 1999) using estimated population percentages from the 2008 NSDUH main study and AMI estimates from the 2008 MHSS clinical study. This allocation minimized the projected SE and RSE of AMI estimates. Because AMI is prevalent in every K6 scoring band, the new allocation increased the sampling rates in the lower K6 ranges and therefore reduced the size of the weights in those K6 groups. Table 3.5 shows the expected sample distribution for the 500 clinical follow-up interviews under this modified design (i.e., based on SMI in quarters 1 and 2 and based on AMI in quarters 3 and 4). The expected overall design effect[8] was 2.0072, which led to an expected effective sample size of 249. The projected SE and RSE of the all-adult estimate of SMI prevalence based on the MHSS clinical sample were 2.39 and 13.92 percent, respectively.

Table 3.5 2009 Mental Health Surveillance Clinical Follow-Up Study Sample Allocation: Modified Design (*n* = 500)

K6 Score	Percent of Population	Assumed AMI Rate (%)	Sample Size
0 to 3	53.10	3.00	96
4 to 5	13.98	13.42	57
6 to 7	9.35	13.95	47
8 to 9	6.08	33.84	59
10 to 11	4.52	43.43	59
12 to 15	6.77	53.78	103
16 or Higher	6.21	76.04	79
Total	**100.00**	**17.15**	**500**

AMI = any mental illness; K6 = Kessler-6, a 6-item psychological distress scale; NSDUH = National Survey on Drug Use and Health.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008.

The probability sample of 500 clinical follow-up interviews was distributed across four calendar quarters with approximately 125 follow-up interviews per quarter. Based on response rate estimates over quarters 1 through 4 of the 2008 MHSS, a 96 percent clinical interview eligibility rate, an 86 percent agreement rate for the clinical follow-up interview, and a 76 percent participation rate among those who agreed to complete the interview were assumed in the sample selection determinations. Table 3.6 displays the 2009 design parameters for quarters 1 through 4. Throughout 2009, the sample was monitored, and the sampling parameters were modified quarterly to ensure that the goal of 500 completions was achieved. In addition, for the last 5 weeks in quarter 4,[9] the probability of selection of the NSDUH respondents for the clinical follow-up survey was set to zero so that cases would not be sampled without adequate time for completion (by December 14, 2009) (see Section 3.8.1 for further details).

[8] The expected design effect for the 2009 MHSS clinical study was the product of the usual design effect for adults in the main survey (about 3.0) and the design effect for the two-phase sample stratified by K6 scores (about 0.7).

[9] The MHSS clinical interview recruitment in 2009 ended on November 16th, the MHSS clinical study data collection ended on December 14th, and the NSDUH main study data collection ended on December 21st.

Table 3.6 Design Parameters for the 2009 Mental Health Surveillance Clinical Follow-Up Study, Quarters 1 through 4

Design Parameter	Quarters 1 and 2 Total	Quarters 3 and 4 Total	Quarters 1 and 2 Expected Persons per Segment	Quarters 3 and 4 Expected Persons per Segment
Interview Respondents Aged 18 or Older, by K6 Score	22,500	22,500	6.250	6.250
Score 0 to 3 (45% of Cases)	10,150	10,150	2.820	2.820
Score 4 to 5 (14% of Cases)	3,205	3,205	0.890	0.890
Score 6 to 7 (10% of Cases)	2,298	2,298	0.638	0.638
Score 8 to 9 (7% of Cases)	1,573	1,573	0.437	0.437
Score 10 to 11 (6% of Cases)	1,248	1,248	0.347	0.347
Score 12 to 15 (9% of Cases)	1,977	1,977	0.549	0.549
Score 16 or Higher (9% of Cases)	1,978	1,978	0.549	0.549
Sampling Rate, by K6 Score[1]				
Score 0 to 3	0.0024	0.0076		
Score 4 to 5	0.0074	0.0138		
Score 6 to 7	0.0130	0.0167		
Score 8 to 9	0.0336	0.0300		
Score 10 to 11	0.0503	0.0379		
Score 12 to 15	0.0589	0.0418		
Score 16 or Higher	0.0448	0.0313		
Selected for Telephone Clinical Follow-Up, by K6 Score				
Score 0 to 3	23	74	0.006	0.020
Score 4 to 5	23	42	0.006	0.012
Score 6 to 7	29	37	0.008	0.010
Score 8 to 9	51	45	0.014	0.013
Score 10 to 11	60	45	0.017	0.013
Score 12 to 15	112	79	0.031	0.022
Score 16 or Higher	85	59	0.024	0.017
Total Selected for Telephone Clinical Follow-Up	382	382	0.106	0.106
Percent Agreeing to Clinical Follow-Up	0.86	0.86		
Percent Completing the Clinical Follow-Up Interview	0.76	0.76		
Completed Clinical Follow-Up Interview, by K6 Score	250	250	0.069	0.069
Score 0 to 3	15	48	0.004	0.013
Score 4 to 5	15	28	0.004	0.008
Score 6 to 7	19	24	0.005	0.007
Score 8 to 9	33	30	0.009	0.008
Score 10 to 11	39	30	0.011	0.008
Score 12 to 15	73	52	0.020	0.014
Score 16 or Higher	56	39	0.015	0.011

K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

[1] The actual sampling rates were slightly higher than those shown in this table. In quarter 1, they were increased to account for the reduced 18 or older sample. In quarter 2, they were higher to compensate for a low clinical interview yield in quarter 1.

NOTE: The estimated distribution of interview respondents is based on the 2008 NSDUH. Estimated MHSS response rates are based on the 2008 Mental Health Surveillance Clinical Follow-Up Study.

Among NSDUH respondents aged 18 or older, 96 percent were eligible to be sampled for the MHSS clinical follow-up study (i.e., they completed the NSDUH main interview in English). At the end of the 2009 data collection, 87 percent of those selected for the follow-up agreed to participate, and 78 percent of those persons completed the clinical interview. The 2009 MHSS resulted in 521 completed clinical interviews. However, one case was excluded from the final dataset because of incomplete data (see Section 3.8.5 for further details). The final number of completed interviews in 2009 was 520. A summary of the 2009 MHSS clinical respondents by quarter is included in Table 3.7.

Table 3.7 2009 Mental Health Surveillance Clinical Follow-Up Study, Quarters 1 through 4 Summary

Design Parameter	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
Interview Respondents Aged 18 or Older	11,403	11,963	11,264	10,979	45,609
Unweighted K6 Distribution					
Score 0 to 3	0.45	0.45	0.45	0.45	0.45
Score 4 to 5	0.14	0.14	0.14	0.14	0.14
Score 6 to 7	0.10	0.10	0.10	0.11	0.10
Score 8 to 9	0.07	0.07	0.07	0.07	0.07
Score 10 to 11	0.06	0.05	0.05	0.06	0.06
Score 12 to 15	0.09	0.09	0.09	0.09	0.09
Score 16 or Higher	0.09	0.09	0.09	0.08	0.09
Eligible for MHSS	10,930	11,452	10,786	10,540	43,708
Eligibility Rate	*0.9585*	*0.9573*	*0.9576*	*0.9600*	*0.9583*
Selected for Telephone Clinical Follow-Up[1]	182	192	211	204	789
Zero Probability Cases	0	0	0	21	21
Agreed to Clinical Follow-Up	156	167	183	159	665
Percent Agreeing to Clinical Follow-Up (Including Zero Probability Cases)	*0.8571*	*0.8698*	*0.8673*	*0.7794*	*0.8428*
Percent Agreeing to Clinical Follow-Up (Excluding Zero Probability Cases)	*0.8571*	*0.8698*	*0.8673*	*0.8689*	*0.8659*
Completed Clinical Follow-Up Interviews	123	125	142	130	520
Clinical Follow-Up Interview Completion Rate	*0.7885*	*0.7485*	*0.7760*	*0.8176*	*0.7820*

K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

[1] Includes cases assigned a zero probability of selection that would have been selected based on their K6 rates.

NOTE: The overall response rate to the MHSS clinical follow-up study should also consider the response rate to the main study. In 2009, the weighted overall response rate for the main NSDUH study was 66.79 percent.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2009.

3.6 2010, 2011, and 2012 Sample Allocations

The 2010 sample was designed to yield 500 clinical interviews, and the 2011 and 2012 samples were designed to yield 1,500 clinical interviews. Because the 2010 through 2012 sample designs were similar, they are all described in this section.

The sample selection algorithms used in 2008 and 2009 led to some respondents having much greater weights than others, which resulted in large SEs of estimates. Because young persons aged 12 to 25 are oversampled in the NSDUH main study, the unweighted sample distribution by age in the main NSDUH sample does not reflect the distribution of the population by age. That is, a respondent aged 18 to 25 in NSDUH represents a smaller proportion of the population than a respondent aged 50 or older; therefore, the younger respondent has a much smaller weight. Because age was not a consideration in the sample selection algorithms in 2008 and 2009, this overrepresentation of young persons was maintained in the clinical sample; thus, the unequal weighting in the main NSDUH sample was maintained in the clinical sample.

Because this oversampling of young adults was not needed for the clinical sample's analytic goals, and the extremely unequal weights were adversely affecting the variance of the

survey estimates, a new selection algorithm was developed for the 2010 through 2012 clinical samples that effectively neutralized the effects of the NSDUH main interview oversampling. Subsamples of eligible respondents were selected for clinical follow-up with probabilities based on their K6 scores and WHODAS scores and on their age. Specifically, an age group equalization factor was used to adjust the K6- and WHODAS-based selection probabilities such that persons were selected for the clinical follow-up sample in accordance with the age distribution in the U.S. adult population. Essentially, this led to young persons in the main NSDUH sample being selected for the clinical follow-up study at a lower probability compared with persons in older age groups. This approach resulted in weights across age groups that were less variable.

Table 3.8 shows some of the age-related factors used to compute sampling rates in 2010, 2011, and 2012. For example, based on the 2008 population estimates and the 2010 planned sample, the average weighting for persons aged 50 or older was almost 10 times as large as the average weight for persons aged 18 to 25. (Smaller differences occurred for the intermediate age groups—26 to 34 and 35 to 49.) To compensate for this initial disparity in weights and to focus on persons aged 18 or older as a whole, sampling rates were set for persons aged 18 to 25, then increased for the other three age groups by adjusting the sampling rate using the final age-related weight equalization factor shown in Table 3.8. For each age group, the derived weight equalization factor is equal to the average weight for that age group divided by the average weight for the 18 to 25 age group. For example, the weight equalization factor in 2010 for the 26 to 34 age group equals 5,656/1,464 ≈ 3.8637 (see Table 3.8).[10] Because the average weight for persons aged 50 or older was so much higher than the other age groups, use of the derived weight equalization factors would have greatly increased the sampling rate for persons aged 50 or older. An adjusted set of factors that partially reduced the unequal weighting effects across age groups was specified instead. Rather than using a different age equalization factor for each age group, the adjusted equalization factors for the 35 to 49 and 50 or older age groups were set equal to the factor for the 26 to 34 age group.

The general sample allocation strategy was to find an allocation that provided a more precise estimate of all-adult SMI prevalence so that appropriate cut points (i.e., points in the SMI predicted probability continuum at which cases would be classified as SMI or not; see Chapter 7 for more details) could be established based on the MHSS clinical sample. To achieve this allocation, a total of 225 strata were defined based on the combination of 25 possible K6 scores (0 to 24)[11] and 9 possible WHODAS scores (0 to 8). The 225 strata were further split based on age group (18 to 25, 26 to 34, 35 to 49, and 50 or older), for a total of 900 allocation cells.

[10] The weight equalization factors were calculated using the average weights with decimals (not rounded average weights).

[11] In the prediction model, a recoded form of K6 score was used: Scores of 0 to 7 were recoded as 0, and all other scores had 7 subtracted from them to give a recoded total ranging from 0 to 17. These scores were reverse recoded to return to the original K6 scores that were used in the two-way matrix. This explains why the predicted probabilities of mental illness are all identical for K6 scores of 0 to 7.

Table 3.8 Age-Related Factors for the 2010 to 2012 Mental Health Surveillance Clinical Follow-Up Study

Year	Age	Population	Interview Respondents Aged 18 or Older	Average Weight	Initial Age-Related Weight Equalization Factor	Eligibility Factor[1] (%)	Response Rate Factor[2] (%)	Final Age-Related Weight Equalization Factor[3]
2010	18 to 25	32,938,184	22,500	1,464	1.0000	96.11	67.43	1.00000
	26 to 34	35,634,108	6,300	5,656	3.8637	93.20	66.87	4.01811
	35 to 49	64,198,531	9,700	6,618	3.8637	94.66	62.33	4.24452
	50 or Older	92,151,942	6,500	14,177	3.8637	96.79	58.72	4.40612
2011	18 to 25	33,579,988	22,500	1,492	1.0000	96.69	69.27	1.00000
	26 to 34	36,214,628	6,000	6,036	4.0442	93.66	62.18	4.65050
	35 to 49	63,166,074	9,000	7,018	4.0442	94.86	67.59	4.22450
	50 or Older	94,245,857	7,500	12,566	4.0442	96.73	65.48	4.27670
2012	18 to 25	34,072,349	22,500	1,514	1.0000	97.08	67.72	1.00000
	26 to 34	36,523,574	6,000	6,087	4.0198	93.03	71.89	3.95170
	35 to 49	62,042,733	9,000	6,894	4.0198	94.64	67.97	4.10880
	50 or Older	96,633,922	7,500	12,885	4.0198	96.76	60.00	4.55250

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

[1] Only NSDUH respondents who had completed the main interview in English were eligible for the clinical follow-up.

[2] Response rates shown are the product of the percentage agreeing to the clinical follow-up survey and the percentage of those who actually participated.

[3] The final age-related equalization factor is the initial weight equalization factor divided by the eligibility and response rate factors and then normalized.

NOTE: Population estimates were made using 2008 NSDUH data for the 2010 design, 2009 NSDUH data for the 2011 design, and 2010 NSDUH data for the 2012 design. Eligibility and response rate factors were computed using 2008 MHSS clinical data for the 2010 design, 2009 MHSS clinical data for the 2011 design, and 2010 MHSS clinical data for the 2012 design.

To determine the sampling rates by stratum (denoted as *h*) and age group, the proportionality factors ($r_{h,age}$) were calculated based on the estimated probability of SMI, the estimated eligibility and response rates, and the final age-related weight equalization factors. Statistical models were developed to estimate the probability of SMI based on the K6 and WHODAS scores, while eligibility rates, response rates, and weight equalization factors were defined based on the age groups. The proportionality factors were defined as follows:

$$r_{h,age} \propto \frac{\sqrt{P_h(1-P_h)}}{E_{age}RR_{age}} * F_{age},$$

where P_h refers to the predicted probability of SMI in stratum *h*, and F_{age}, E_{age}, and RR_{age} refer to the age-specific weight equalization factors, eligibility factors, and response rate factors, respectively. These proportionality factors then were multiplied by the projected sample counts and scaled to achieve the desired overall respondent sample (500 persons aged 18 or older in 2010 and 1,500 persons aged 18 or older in 2011 and 2012) and to obtain the stratum and age-specific sampling rates.

As an example, from the 2011 MHSS clinical data the predicted probability of SMI for a person with a K6 score of 10 and a WHODAS score of 6 was 0.1398. For the 18 to 25 age group, the proportionality factor then would be

$$r_{h,18-25} = \frac{\sqrt{0.1398(1-0.1398)}}{0.9669*0.6927} *1.000 = 0.5177.$$

The proportionality factors were then scaled in order to achieve an overall sample of 1,500 persons. A scaling factor of 0.0885 was expected to yield 1,500 completed interviews, so it was applied to each proportionality factor. Thus, the sampling rate for this stratum and age group was 0.5177 * 0.0885 = 0.0458.

Projected yields of positive cases based on the predicted probability of SMI and AMI broken out by age group and year are provided in Table 3.9. Table 6.1 in Chapter 6 presents the total number of Structured Clinical Interview for DSM-IV (SCID) respondents for each age group by year,[12] and Table 6.2 presents the total number of SCID respondents who were diagnosed with SMI or AMI by year in comparison.

In addition, Table 3.10 provides the 2010-2012 MHSS clinical sample allocations by K6 group, and Table 3.11 provides the 2010-2012 MHSS clinical sample allocations by WHODAS score. The 2010 probability sample of 500 clinical follow-up interviews was distributed across four calendar quarters with approximately 125 clinical follow-up interviews per quarter. The 2011 and 2012 samples of 1,500 clinical follow-up interviews were distributed across four calendar quarters with approximately 375 clinical follow-up interviews per quarter. Throughout the 2010, 2011, and 2012 surveys, the MHSS clinical sample was monitored, and the sampling

[12] See Chapter 1 for details on the Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP or SCID).

Table 3.9 Projected Yields of Predicted Positive Cases, by Age Group: 2010 to 2012 MHSS Clinical Follow-Up

Year	Measure	18 to 25	26 to 34	35 to 49	50 or Older	18 or Older
2010	SMI	25	23	31	10	89
	AMI	67	59	80	33	239
	Total	**116**	**116**	**170**	**98**	**500**
2011	SMI	72	71	92	39	274
	AMI	194	184	226	118	721
	Total	**335**	**343**	**477**	**345**	**1,500**
2012	SMI	76	69	92	45	281
	AMI	195	174	224	126	720
	Total	**338**	**336**	**476**	**351**	**1,500**

AMI = any mental illness; K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: Predicted probabilities of SMI and AMI were calculated for each of the 225 K6 by WHODAS strata using models based on 2008 Mental Health Surveillance Clinical Follow-Up Study data. These probabilities were applied to the anticipated MHSS Clinical Follow-Up Study distributions for 2010 through 2012 to obtain estimates at the age group level. Sample distributions were estimated using 2008 NSDUH data for the 2010 design, 2009 NSDUH data for the 2011 design, and 2010 NSDUH data for the 2012 design.

Table 3.10 2010 to 2012 Mental Health Surveillance Clinical Follow-Up Study Sample Allocations, by K6 Group

Year	K6 Group	Percent of Population[1]	Assumed SMI Rate (%)[2,3]	Expected Sample Size	Expected SMI Count	Overall Sampling Rate
2010	0 to 3	53.28	0.91	162	0	0.00794
	4 to 5	14.22	1.20	51	0	0.00790
	6 to 7	9.28	1.73	38	0	0.00841
	8 to 9	6.06	2.98	34	0	0.01060
	10 to 11	4.74	5.01	35	1	0.01373
	12 to 15	6.55	12.65	78	18	0.01958
	16 or Higher	5.87	39.03	103	69	0.02620
	Total	**100.00**	**4.36**	**500**	**89**	
2011	0 to 3	53.21	0.92	491	0	0.02322
	4 to 5	13.78	1.21	148	0	0.02293
	6 to 7	9.41	1.61	116	0	0.02418
	8 to 9	5.96	2.65	96	0	0.03034
	10 to 11	4.64	5.32	101	3	0.03978
	12 to 15	6.99	12.49	233	50	0.05569
	16 or Higher	6.02	41.07	316	221	0.07751
	Total	**100.00**	**4.56**	**1,500**	**274**	
2012	0 to 3	53.54	0.92	497	0	0.02370
	4 to 5	13.67	1.21	145	0	0.02386
	6 to 7	9.08	1.58	114	0	0.02534
	8 to 9	6.26	2.86	98	0	0.03126
	10 to 11	4.41	5.51	99	4	0.04192
	12 to 15	6.71	12.28	225	52	0.05731
	16 or Higher	6.33	41.85	322	225	0.07980
	Total	**100.00**	**4.70**	**1,500**	**281**	

K6 = Kessler-6, a 6-item psychological distress scale; NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

[1] Source: 2008 NSDUH data for the 2010 design, 2009 NSDUH data for the 2011 design, and 2010 NSDUH data for the 2012 design.

[2] Source: 2008 NSDUH clinical follow-up study for the 2010 design, 2009 NSDUH clinical follow-up study for the 2011 design, and 2010 NSDUH clinical follow-up study for the 2012 design.

[3] To compute assumed SMI rates, SMI estimates by K6 and WHODAS score were averaged (weighted) across K6 scores. These rates are not the actual SMI rates that were used in the sample allocation.

Table 3.11 2010 to 2012 Mental Health Surveillance Clinical Follow-Up Study Sample Allocations, by WHODAS Score

Year	WHODAS Score	Percent of Population[1]	Assumed SMI Rate (%)[2,3]	Expected Sample Size	Expected SMI Count	Overall Sampling Rate
2010	0	74.80	1.09	248	0	0.00736
	1	7.11	2.71	39	1	0.01207
	2	5.30	4.26	36	1	0.01498
	3	3.14	8.01	32	2	0.02276
	4	2.57	11.80	29	7	0.02492
	5	2.04	19.04	29	11	0.03140
	6	1.74	31.80	31	21	0.04003
	7	1.32	41.36	25	18	0.04255
	8	1.98	54.53	31	28	0.03518
	Total	**100.00**	**4.36**	**500**	**89**	
2011	0	74.50	1.05	734	0	0.02188
	1	7.21	2.49	115	2	0.03529
	2	4.99	4.55	108	5	0.04829
	3	3.32	8.51	97	9	0.06485
	4	2.48	13.90	92	23	0.08205
	5	2.38	20.42	95	38	0.08876
	6	1.93	30.34	90	56	0.10363
	7	1.28	43.47	70	55	0.12286
	8	1.92	58.32	100	87	0.11535
	Total	**100.00**	**4.56**	**1,500**	**274**	
2012	0	74.26	1.06	738	0	0.02210
	1	7.41	2.61	116	3	0.03473
	2	4.95	4.67	104	4	0.04695
	3	3.37	7.68	93	8	0.06132
	4	2.61	13.45	89	22	0.07585
	5	2.05	21.26	89	36	0.09634
	6	1.83	32.44	87	57	0.10606
	7	1.44	42.75	75	56	0.11673
	8	2.10	58.81	108	95	0.11413
	Total	**100.00**	**4.70**	**1,500**	**281**	

K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Scale.

[1] Source: 2008 NSDUH data for the 2010 design, 2009 NSDUH data for the 2011 design, and 2010 NSDUH data for the 2012 design.

[2] Source: 2008 MHSS clinical follow-up study for the 2010 design, 2009 MHSS clinical follow-up study for the 2011 design, and 2010 MHSS clinical follow-up study for the 2012 design.

[3] To compute assumed SMI rates, SMI estimates by K6 and WHODAS score were averaged (weighted) across WHODAS scores. These rates are not the actual SMI rates that were used in the sample allocation.

parameters were modified on an as-needed basis. Sampling rates were adjusted as needed to ensure that the targeted number of completions was achieved for each year. For the 2010 and 2011 clinical study, for the last 3 weeks in quarter 4,[13] the probability of selection of the NSDUH respondents for the clinical follow-up survey was set to zero (see Section 3.8.1 for further details). For the 2012 MHSS, the probability of selection of the NSDUH respondents for the clinical follow-up survey was not set to zero. Respondents were selected for clinical follow-up through December 20, 2012 (the last day of NSDUH main study data collection), and MHSS clinical data collection continued through January 17, 2013.

The 2010 MHSS clinical study resulted in 516 completed clinical interviews. Approximately 84 percent of selected persons agreed to participate, and 81 percent of those persons completed the MHSS clinical interview. A summary of the 2010 respondents by quarter is included in Table 3.12.

The 2011 MHSS clinical study resulted in 1,495 completed clinical interviews. An estimated 84 percent of selected persons agreed to participate, and 79 percent of those persons completed the MHSS clinical interview. A summary of the 2011 respondents by quarter is included in Table 3.13.

The 2012 MHSS clinical study resulted in 1,622 completed clinical interviews. An estimated 84 percent of selected persons agreed to participate, and 79 percent of those persons completed the MHSS clinical interview. A summary of the 2012 respondents by quarter is included in Table 3.14.

3.7 Creation of Variance Estimation Strata and Replicates

The nature of the NSDUH stratified, clustered sampling design requires that the design structure be taken into consideration when computing variances of survey estimates. Variance estimation strata and replicates specific to the clinical follow-up study were created to appropriately account for the sample design. To define the variance estimation strata for the MHSS clinical study, the 900 NSDUH main study variance strata were sorted, and groups of 9 adjacent main study variance strata were collapsed for each year of data collection. Thus, a total of 100 MHSS clinical interview variance strata were formed (and named MHVESTR). The number of variance strata formed was selected to optimize the number of degrees of freedom (*df*) available for analyses. The variance strata were formed for the purpose of producing national estimates only because subnational estimates are not recommended. It is recommended that the variable MHVESTR be used to produce SEs of estimates when using the combined 2008-2012 clinical data.

In this report, methodological analyses of the clinical interview data by individual year are included, even though the clinical interview data are intended to be used as a 5-year pooled dataset when conducting nonmethodological, substantive analyses. Prior to the completion of data collection, a preliminary strata variable (MHVSTR09) with only 50 strata was created and

[13] The MHSS clinical interview recruitment in both 2010 and 2011 ended on November 29th, and the MHSS clinical study and the NSDUH main study data collection ended on December 20th.

Table 3.12 2010 Mental Health Surveillance Clinical Follow-Up Study, Quarters 1 through 4 Summary

Design Parameter	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
Interview Respondents Aged 18 or Older	10,877	12,102	11,844	11,021	45,844
Unweighted K6 Distribution, by K6 Score					
Score 0 to 3	0.45	0.46	0.47	0.45	0.46
Score 4 to 5	0.14	0.14	0.14	0.13	0.14
Score 6 to 7	0.10	0.10	0.10	0.11	0.10
Score 8 to 9	0.07	0.07	0.07	0.07	0.07
Score 10 to 11	0.06	0.05	0.05	0.05	0.05
Score 12 to 15	0.09	0.08	0.09	0.09	0.09
Score 16 or Higher	0.09	0.09	0.09	0.09	0.09
Eligible for MHSS	10,446	11,608	11,341	10,563	43,958
Eligibility Rate	*0.9604*	*0.9592*	*0.9575*	*0.9584*	*0.9589*
Selected for Telephone Clinical Follow-Up[1]	190	246	175	157	768
Zero Probability Cases	0	0	0	4	4
Agreed to Clinical Follow-Up	163	198	146	133	640
Percent Agreeing to Clinical Follow-Up (Including Zero Probability Cases)	*0.8579*	*0.8049*	*0.8343*	*0.8471*	*0.8333*
Percent Agreeing to Clinical Follow-Up (Excluding Zero Probability Cases)	*0.8579*	*0.8049*	*0.8343*	*0.8693*	*0.8377*
Completed Clinical Follow-Up Interviews	132	157	115	112	516
Clinical Follow-Up Interview Completion Rate	*0.8098*	*0.7929*	*0.7877*	*0.8421*	*0.8063*

K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; WHODAS = World Health Organization Disability Assessment Scale.

[1] Includes cases assigned a zero probability of selection that would have been selected based on their K6 and WHODAS scores.

NOTE: The overall response rate to the MHSS clinical follow-up study should also consider the response rate to the main study. In 2010, the weighted overall response rate for the main NSDUH study was 65.94 percent.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2010.

used for the initial modeling and methodological work that analyzed data from each individual year. Because of the smaller yearly sample sizes, larger clinical interview variance strata were needed. Therefore, analyses in this report also used that preliminary version.

Two replicates (i.e., primary sampling units for variance estimation purposes) are defined within each NSDUH main study variance stratum. Each variance replicate consists of four segments, one for each quarter of data collection. The first replicate consists of those segments that are "phasing out" or will not be used in the next survey year. The second replicate consists of those segments that are "phasing in" or will be fielded again the following year, thus constituting the 50 percent overlap between survey years. The variance replicate assigned to the segment in which the MHSS clinical interview respondent was sampled was retained for the MHSS clinical study (and named MHVEREP). Further details about the formation of NSDUH main study

Table 3.13 2011 Mental Health Surveillance Clinical Follow-Up Study, Quarters 1 through 4 Summary

Design Parameter	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
Interview Respondents Aged 18 or Older	10,840	12,481	12,170	11,108	46,599
Unweighted K6 Distribution, by K6 Score					
Score 0 to 3	0.46	0.47	0.46	0.45	0.46
Score 4 to 5	0.14	0.13	0.13	0.14	0.14
Score 6 to 7	0.10	0.10	0.10	0.10	0.10
Score 8 to 9	0.07	0.07	0.07	0.07	0.07
Score 10 to 11	0.05	0.05	0.06	0.05	0.05
Score 12 to 15	0.09	0.08	0.09	0.09	0.09
Score 16 or Higher	0.09	0.09	0.09	0.09	0.09
Eligible for MHSS	10,392	11,974	11,665	10,709	44,740
Eligibility Rate	*0.9587*	*0.9594*	*0.9585*	*0.9641*	*0.9601*
Selected for Telephone Clinical Follow-Up[1]	543	672	531	531	2,277
Zero Probability Cases	15	0	0	26	41
Agreed to Clinical Follow-Up	450	561	449	421	1,881
Percent Agreeing to Clinical Follow-Up (Including Zero Probability Cases)	*0.8287*	*0.8348*	*0.8456*	*0.7928*	*0.8261*
Percent Agreeing to Clinical Follow-Up (Excluding Zero Probability Cases)	*0.8523*	*0.8348*	*0.8456*	*0.8337*	*0.8412*
Completed Clinical Follow-Up Interviews	363	436	359	337	1,495
Clinical Follow-Up Interview Completion Rate	*0.8067*	*0.7772*	*0.7996*	*0.8005*	*0.7948*

K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; WHODAS = World Health Organization Disability Assessment Scale.

[1] Includes cases assigned a zero probability of selection that would have been selected based on their K6 and WHODAS scores.

NOTE: The overall response rate to the MHSS clinical follow-up study should also consider the response rate to the main study. In 2011, the weighted overall response rate for the main NSDUH study was 64.69 percent.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2011.

variance strata and replicates can be found in the 2012 sample design report (Morton, Martin, Shook-Sa, Chromy, & Hirsch, 2013). When analyzing data from the MHSS clinical study, the variance estimation strata, replicates, and analysis weight must be taken into account using software capable of design-consistent estimation (e.g., SUDAAN®, STATA®).[14]

[14] SUDAAN® is used on NSDUH and stands for SUDAAN® Software for Statistical Analysis of Correlated Data (RTI International, 2012). STATA® is an integrated statistical software package that can be used for data analysis, data management, and graphics (http://www.stata.com/).

Table 3.14 2012 Mental Health Surveillance Clinical Follow-Up Study, Quarters 1 through 4 Summary

Design Parameter	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Total
Interview Respondents Aged 18 or Older	10,894	12,144	11,997	10,801	45,836
Unweighted K6 Distribution, by K6 Score					
Score 0 to 3	0.45	0.46	0.46	0.46	0.46
Score 4 to 5	0.14	0.14	0.14	0.14	0.14
Score 6 to 7	0.10	0.10	0.10	0.10	0.10
Score 8 to 9	0.07	0.06	0.07	0.07	0.07
Score 10 to 11	0.05	0.06	0.06	0.05	0.06
Score 12 to 15	0.09	0.09	0.09	0.08	0.09
Score 16 or Higher	0.09	0.09	0.09	0.09	0.09
Eligible for MHSS	10,488	11,699	11,591	10,442	44,220
Eligibility Rate	*0.9627*	*0.9634*	*0.9662*	*0.9668*	*0.9647*
Selected for Telephone Clinical Follow-Up[1]	585	596	634	649	2,464
Zero Probability Cases	0	0	0	0	0
Agreed to Clinical Follow-Up	491	497	532	543	2,063
Percent Agreeing to Clinical Follow-Up (Including Zero Probability Cases)	*0.8393*	*0.8339*	*0.8391*	*0.8367*	*0.8373*
Percent Agreeing to Clinical Follow-Up (Excluding Zero Probability Cases)	*0.8393*	*0.8339*	*0.8391*	*0.8367*	*0.8373*
Completed Clinical Follow-Up Interviews	384	379	434	425	1,622
Clinical Follow-Up Interview Completion Rate	*0.7821*	*0.7626*	*0.8158*	*0.7827*	*0.7862*

K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; WHODAS = World Health Organization Disability Assessment Scale.

[1] Includes cases assigned a zero probability of selection that would have been selected based on their K6 and WHODAS scores.

NOTE: The overall response rate to the MHSS clinical follow-up study should also consider the response rate to the main study. In 2012, the weighted overall response rate for the main NSDUH study was 62.87 percent.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2012.

3.8 Issues That Affected the Clinical Follow-Up Sample Design

3.8.1 Assigning NSDUH Clinical Interview-Eligible Cases Zero Selection Probability

The clinical interviews were conducted by telephone follow-up after the completion of the main NSDUH interview, so it was a logistical challenge for RTI to collect data from respondents selected for the MHSS clinical study late in quarter 4 because the NSDUH protocol is to suspend all data collection activities approximately 10 days prior to December 31st of the year. From 2008 to 2011, recruitment for the clinical interview was suspended early in order to complete data collection in time. This suspension of recruitment was accomplished by altering the clinical interview sampling algorithm in the main NSDUH interview such that respondents

who completed their main study interviews beyond the specified clinical interview cutoff date[15] would have a zero probability of selection for clinical follow-up. Respondents who completed their main NSDUH interview after the clinical interview cutoff point had no chance of being selected for the MHSS clinical interview, regardless of their K6 scores, WHODAS scores, or age. This suspension of recruitment for the clinical interview was not implemented in the 2012 MHSS. In 2012, recruitment for the clinical interviews continued through the end of main study data collection,[16] and selected cases continued to be contacted and interviewed after main study data collection concluded.

In addition to the cases that had no chance of selection because their main study interview was conducted after the clinical interview cutoff date, at the beginning of quarter 1 in the 2011 data collection, 15 NSDUH respondents who should have been selected for the clinical interview were inadvertently given a zero probability of selection for the MHSS clinical interview.

Eligible respondents who were given a zero probability of selection for the MHSS clinical interview study may have different mental health characteristics than persons who were given a chance of being selected. To avoid potential bias from the exclusion of zero probability cases, cases were identified that would have been selected for the MHSS clinical study based on their age (for 2010 to 2012 only), K6 scores, and WHODAS scores, and these zero probability cases were classified as nonrespondents rather than treated as not sampled cases. In 2008 and 2009, MHSS clinical interview selection status for the zero probability cases was determined by comparing the sampling rates for the cases' K6 rates with their random numbers to determine whether they would have been selected. For 2010 and 2011, the probability of selection took into account their K6 rates, WHODAS scores, and age group adjustments. All zero probability cases that would have been selected if they had not been given a zero probability of selection were treated as nonrespondents in the calculation of the MHSS clinical sample analysis weights (see Section 5.4 in Chapter 5 for more details on weighting adjustments for nonresponse).
The numbers of cases assigned a zero probability of selection that would otherwise have been selected for the clinical interview are shown in Tables 3.4, 3.7, 3.12, 3.13, and 3.14 for each year of MHSS clinical study data collection.

3.8.2 Noncoverage for Non-English Speakers

The target population for the MHSS clinical study excluded persons whose main study NSDUH interview was conducted in Spanish. Approximately 4 percent of the main NSDUH interviews are completed in Spanish each year. Among interviews with Hispanic respondents, 23 percent are conducted in Spanish.[17] The undercoverage of Spanish speakers was adjusted through a sample weighting procedure (see Section 5.3 in Chapter 5).

[15] Recruitment for the clinical interview was suspended with 4 weeks remaining in data collection in 2008 (November 21, 2008), with 5 weeks remaining in data collection in 2009 (November 16, 2009), and with 3 weeks remaining in data collection in 2010 and 2011 (November 29, 2010, and November 29, 2011).

[16] Both the main study data collection and the MHSS clinical interview recruitment were suspended on December 20, 2012. Clinical interviews were conducted through January 17, 2013.

[17] This percentage is based on 2012 NSDUH respondents aged 18 or older.

3.8.3 Gulf Coast Oversample

In 2011, a Gulf Coast Oversample (GCO) was included to measure the impact of the April 20, 2010, Deepwater Horizon oil spill on substance use, mental health, and the utilization of substance abuse and mental health services in the Gulf Coast region's areas of Louisiana, Mississippi, Alabama, and Florida. The 2011 main study sample was expanded by 2,000 completed interviews in specific counties and/or parishes in Alabama, Florida, Louisiana, and Mississippi.

With the additional sample in the Gulf Coast region, the main study sample shifted from approximately 1.3 percent in the affected area to 3.3 percent in this area. As a result, 56 MHSS clinical interviews (3.75 percent of the total) were completed in the affected area. The number of completed main study and clinical interviews in the Gulf Coast region by year are presented in Table 3.15.

Table 3.15 Completed NSDUH Main Interviews and MHSS Clinical Interviews in the GCO Region, by Year

Year	NSDUH Main Interviews	MHSS Clinical Interviews
2008	753	15
2009	783	5
2010	802	7
2011	2,313	56
2012	803	22
Total	**5,454**	**105**

GCO = Gulf Coast Oversample; MHSS=Mental Health Surveillance Study; NSDUH=National Survey on Drug Use and Health.

NOTE: The GCO region comprises the geographic areas in the GCO. The Gulf Coast Oil Spill Disaster Area includes the 32 counties and parishes in Louisiana, Mississippi, Alabama, and Florida that were most affected by the April 2010 oil spill: Iberia, Jefferson, Lafayette, Lafourche, Orleans, Plaquemines, St. Bernard, St. Martin, St. Mary, St. Tammany, Terrebonne, and Vermilion of Louisiana; George, Hancock, Harrison, Jackson, Pearl River, and Stone of Mississippi; Baldwin, Clarke, Escambia, Mobile, Monroe, and Washington of Alabama; and Bay, Escambia, Franklin, Gulf, Okaloosa, Santa Rosa, Wakulla, and Walton of Florida.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

3.8.4 Falsified Main NSDUH Interviews

At the beginning of quarter 4 of the 2011 NSDUH, it was discovered that an FI in Pennsylvania had been falsifying interviews throughout 2011. Data monitoring and field verification are part of a routine process to ensure the quality of NSDUH data. In a typical quarter when only a few falsified cases are identified, the falsified screening and interview cases are reworked, and the responsible FI is removed from the project. Screening and interview cases completed in earlier quarters by FIs found to have falsified data in the current quarter are typically not removed or reworked. Because of the large-scale falsification found in Pennsylvania, the procedure was changed to discard falsified data back to 2008 (when this FI began falsifying clinical records) in addition to reworking the quarter 4 cases.

All main study NSDUH screening interviews completed by this FI from 2008 through 2011 were deemed incomplete with unknown eligibility, and the corresponding interviews were dropped from the analysis files (Morton et al., 2012). Consequently, all clinical follow-up cases sampled from these falsified main study NSDUH interviews were recoded as main study incompletes. MHSS clinical data were reprocessed such that cases selected for the MHSS clinical interviews that were sampled from the falsified main study NSDUH interviews were no longer treated as sampled for the MHSS clinical interviews. MHSS clinical analysis weights were recalculated for 2008 through 2010 without these cases (i.e., the falsified cases were excluded from the nonresponse models and were considered as being neither respondents nor nonrespondents).

As shown in Table 3.16, the removal of falsified cases resulted in two MHSS clinical interviews being removed from the 2008 MHSS clinical follow-up. No completed clinical interviews were lost from the 2009 and 2010 MHSS clinical follow-ups, and the falsification was discovered prior to the processing of the 2011 MHSS clinical follow-up, so no reweighting was required.

Table 3.16 Falsified Cases Removed from the 2008-2010 Mental Health Surveillance Clinical Follow-Up Study

Type of Case Removed	2008 MHSS	2009 MHSS	2010 MHSS
Selected for Clinical Follow-Up	7	3	1
Agreed to Clinical Follow-Up	4	0	0
Completed Clinical Interviews	2	0	0

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH MHSS clinical sample, 2008-2012.

3.8.5 Respondents Excluded from 2008-2012 Adult Clinical Interview Data File

In 2008 and 2009, data from five respondents were excluded and treated as nonrespondents (see Section 5.4 in Chapter 5 for details on the nonresponse weighting adjustment for the SCID sample) in the final SCID dataset. These respondents were excluded either because the respondent had an extremely large weight or because responses on all K6 and WHODAS (or SDS) item scores were missing.

Respondents with low K6 total scores typically had relatively large weights because they were undersampled for the clinical follow-up. One case belonging to the 2008 WHODAS half sample with a large weight was designated as SMI positive by the SCID interview,[18] but the K6 and WHODAS total scores were zero, thus ensuring that this case would always be a false negative in the receiver operating characteristic (ROC) modeling analyses. The large weight (1.6 million for this respondent) had the effect of unduly influencing the ROC models developed in 2008, so this respondent was dropped from the dataset. Two respondents belonging to the 2008 SDS half sample also were removed for similar reasons (i.e., these cases had large weights

[18] See footnote 12.

and were designated as SMI positive by the SCID interview, which unduly influenced the ROC models). After the first two quarters of the data collection, a decision was made to allocate the clinical sample based on AMI rather than SMI in order to reduce the selection probabilities of respondents with extremely large weights while maintaining the efficiency of estimating SMI and AMI directly based on the clinical sample (see Section 3.5 for more information). In 2010 through 2012, in order to decrease the variability of the weights, the sample selection algorithm for the clinical sample was changed further so that the probability of selection was based on the respondent's age, K6 score, and WHODAS score. The new algorithm effectively neutralized the effects of the NSDUH main interview oversampling of young adults (see Section 3.6 for more information).

An additional 2008 respondent from the SDS half sample and a respondent from the 2009 MHSS clinical sample were removed because all of their item scores for the K6 components were missing. For respondents whose item scores for K6 and WHODAS (or SDS) were partially missing, a zero-assignment imputation method was used for both the item scores and total scores. Section 4.3 in Chapter 4 provides more details on how the missing values in the model predictors were treated.

4. Data Collection and Processing

4.1 Background

As stated in Chapter 1, the primary purpose of the Mental Health Surveillance Study (MHSS) is to provide clinical data for estimating a statistical model that predicts serious mental illness (SMI) based on data collected through the National Survey on Drug Use and Health (NSDUH). In Chapter 7, details are given on how this statistical model is developed and utilized for estimating SMI and other levels of mental illness. This chapter describes the data collection and process procedures used for the MHSS clinical study and the imputation procedures used on the NSDUH main study variables that are employed as predictors in the SMI prediction model.

4.2 SCID Data Collection

As discussed in Chapter 3, MHSS clinical interview respondents were selected and recruited for the follow-up clinical interview at the end of the NSDUH main study. Respondents agreeing to participate were asked to provide contact information, including first name, telephone number, an alternate telephone number if available, and the best days and times to call. The field interviewer (FI) entered the respondent's contact information into the NSDUH laptop and transmitted the data to RTI International. The next step was for a data collection manager to review the case and assign it to a clinical interviewer. The data collection manager took into consideration the respondent's time zone and best times and dates for contact, as well as the clinical interviewer's availability and time zone when assigning cases. The clinical interviewer made the first attempt to contact a respondent within 24 hours of receiving the assigned case to schedule an appointment for the interview. If the respondent was not reached on the first call attempt, subsequent calls were made during the respondent's preferred time frame. After the clinical interviewer made several attempts without success during the time period specified by the respondent, calls were made at times outside of the respondent's preferred time frame in an attempt to reach the respondent.

The clinical interviewer would contact the respondent by telephone to schedule and to conduct the interview. Clinical interviewers were instructed to verify that the correct respondent had been contacted, obtain informed consent, and obtain permission to record the interview using the script provided. Permission to record the interview was not a requirement to participate in the study; however, clinical interviewers were instructed to inquire about any refusal to record the interview and to address any respondent concerns, if possible. To ensure confidentiality and privacy, no identifying information was written in the clinical interview booklet or solicited during the recorded interview.

As discussed in Section 2.3 in Chapter 2, the follow-up interview was a modified version of the Structured Clinical Interview for the DSM-IV, or SCID (First et al., 2002). The SCID is a semistructured diagnostic interview used to assess psychiatric disorders according to the criteria in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV; American Psychiatric Association, 1994). As a semistructured clinical interview, the SCID contains structured, standardized questions that are read verbatim and sequentially, then

combined with unstructured follow-up questions that the clinical interviewer tailors to the respondent based on clinical judgment and respondent reports. The SCID was administered over the telephone by clinical interviewers who had undergone extensive training with SCID experts from RTI and the SCID's developer from Columbia University.

Clinical interviewers were trained to use their clinical judgment to code each item based on the respondent's answers and presentation. Using the SCID booklet, each criterion symptom was coded as "1" (absent or false), "2" (subthreshold), "3" (threshold or true), or "?" (inadequate information). Each diagnosis assessed was coded as "1" (absent) or "3" (present). As discussed in Section 4.3.1, the clinical interviewer would also score the Global Assessment of Functioning (GAF) on a scale of 1 (persistent danger to self or others) to 100 (superior functioning, no symptoms). The GAF is based on the respondent's psychological, social, and occupational functioning. Clinical interviewers scored the GAF based on the respondent's worst functioning in the past 12 months. At the conclusion of the SCID, the clinical interviewer read the end of interview script, which thanked the respondent for his or her time and mentioned the possibility of needing to speak with a counselor. Clinical interviewers ensured that the respondent had the toll-free number to the National Lifeline Network should he or she wanted additional information about mental health services in his or her area. If the clinical interviewer felt that the respondent was at all in distress or potentially was a danger to himself or herself or others, a distressed respondent protocol (DRP) was followed.

In order to ensure that both the main NSDUH interview and the clinical interview covered the same reference period, each follow-up interview was completed over the telephone within 4 weeks of the date of the main NSDUH interview. After the call ended, the clinical interviewer completed the interviewer debriefing questions within the SCID and reviewed the booklet to ensure that the documentation was complete. As part of the interviewer debriefing section, clinical interviewers recorded information about problems encountered during the interview (i.e., distressed respondent, cognitive impairment, problems with privacy, comprehension, cooperation), stressors the respondents had experienced in the past 12 months, a diagnosis that needed further assessment, and the overall validity of the SCID data. If clinical interviewers were uncertain how to code a particular item or were feeling unsure about whether they had gathered enough information, they were instructed to consult with one of the clinical supervisors for his or her opinion within an hour of completing the interview. Within 48 hours after completion of the interview, the clinical interviewer uploaded the recorded audio file to the Web-based Case Management System (CMS) and edited and shipped the paper SCID to RTI for final editing and keying in-house. Once shipped, the clinical interviewer assigned the case a code of "80: Materials Shipment (SCID)" and entered the Federal Express tracking number in the notes. A data collection manager then tracked the shipment to ensure that the SCID arrived at RTI.

4.3 SCID Data Processing

When the SCID was received at RTI, it was assigned a status code of "81: SCID Received at Regent." The SCID was then reviewed by clinical supervisors, technical editors, and keying clerks, following the editing and keying process described the following sections.

4.3.1 Clinical Editing

When the SCID was delivered to the MHSS clinical supervisors, it was assigned a status code of "82: SCID Received at Editing." The clinical supervisors reviewed the data collected in all SCID booklets, item by item, comparing the notes provided by the clinical interviewer and the diagnostic rating and listening to the accompanying audio files as needed to ensure confidence in the data. The audio file (if recorded) was matched with the corresponding paper-and-pencil SCID booklet. For full reviews, both the audio recordings and the SCID booklet were reviewed in their entirety. Newly hired clinical interviewers received full reviews of their first two interviews until they demonstrated proficiency in administering all SCID modules. Of the remaining interviews, 10 percent were randomly selected for a full review of the audio recordings and SCID booklet in their entirety. For the remaining partial reviews, significant portions of the overview and selected modules were reviewed. Cases that presented a more complex clinical picture received a closer review (i.e., cases in which the constellation of symptoms, the clinical interviewer's notes, and/or what is known about the indicated disorder[s] warranted special attention, as determined by the clinical supervisor).

As part of the review, clinical supervisors ensured that the clinical interviewers followed project protocol for administering the interview and provided accurate and sufficient notes in the SCID booklet. The clinical supervisors also evaluated the clinical interviewers' clinical interviewing techniques, diagnostic skills, and the extent to which they captured the overall clinical picture relative to the symptoms, diagnoses, and GAF score. In the back of the SCID booklet, clinical supervisors rated the overall validity of the data and noted if there were any other disorders that needed further assessment, such as a disorder that was not assessed in the study or a disorder where more information was needed for diagnosis.

Quality control concerns arose when either a respondent did not provide complete and accurate information or the clinical interviewer did not adequately conceptualize the case or assess the symptoms. To ensure quality control throughout the clinical editing process, the clinical supervisors met to discuss complex cases and data that were difficult to code. Consensus ratings were used to code these data.

For each interview, a clinical supervisor completed a clinical editing form, noting strengths and areas for improvement for the clinical interviewer. The SCID editing forms provided a record of clinical editing, a format for individualized feedback, and a system for monitoring clinical interviewer performance. Completed SCID editing forms were filed in separate folders for each clinical interviewer in a locked filing cabinet. Once clinical editing was completed, the corresponding code of "88: SCID Delivered to Technical Editing" was entered into the CMS, and the SCID booklet was given to the technical editors.

Completed cases that one clinical supervisor considered to be of questionable validity were subject to an independent second review by at least one other clinical supervisor. To address the potential bias to exclude more complex or unusual cases and those with a difficult respondent, the level of difficulty and overall confidence in the case were considered prior to making a final disposition. If all of the clinical supervisors agreed, then the case was removed from the study. In these situations, a code of "90: SCID Not Keyed - Reason Unspecified"

(prior to 2010) or "90.X" (after 2010; "X" depending on the reason) was entered into the CMS. Any relevant feedback was given directly to the clinical interviewer.

4.3.2 Technical Editing

Once clinical editing was completed, technical editors reviewed each SCID to ensure that all variables met the criteria for consistency[19] and to verify that skip logic criteria were true. For example, both variables S1 and E28 in the SCID are used to code the same symptom (any panic attacks); therefore, consistency checks were done such that "IF S1 = 1, then E28 = 1" and to verify that conditions stated in the "skip logic" criteria were true. An example of the verification of the skip logic includes that for major depressive episode (MDE). If criterion A for past year MDE is not met and therefore absent, the remaining items to assess past year MDE are "skipped" and left blank, such that "IF A10 = 1, then A11-A17 = BLANK" (see the technical editing guide in Appendix H of the 2008-2012 MHSS operations report; CBHSQ, 2014a). Any inconsistencies or cases in which the conditions were not true were noted on a technical editing form. The technical editing form and its respective SCID were returned to the clinical supervisors for final review and correction. Once completed by the technical editors, a code of "89: Technical Editing Complete" was recorded in the CMS.

4.3.3 Keying

After technical editing was completed and any discrepancies were resolved, the data clerks keyed the SCID data into a computer and stored the keyed data in an electronic file. All codes that did not meet specifications were flagged by the machine edit program. A report of flagged cases was sent to the clinical supervisors indicating the line(s) where the errors were suspected. SCID booklets flagged with potential problems were retrieved from storage and compared with the computer-generated report. Those line(s) were reviewed again to ensure accuracy in both clinical and technical editing. Any errors were corrected and communicated to the dataset manager to be reset by hand if needed.

4.4 Missing Data

To ensure that complete cases were available for modeling mental illness, all missing values for item scores from the Kessler-6 (K6), World Health Organization Disability Assessment Schedule (WHODAS), and Sheehan Disability Scale (SDS) sets of variables were recoded with zeros, and the recoded item scores were used to calculate the total scores that were used as predictors in the SMI models (see Chapter 7 for more information on the SMI modeling). The item score variables that were recoded include the following:

- past month K6 variables (DSTNRV30, DSTHOP30, DSTRST30, DSTCHR30, DSTEFF30, DSTNGD30),
- past year K6 variables (DSTNRV12, DSTHOP12, DSTRST12, DSTCHR12, DSTEFF12, DSTNGD12), and

[19] The modified SCID questionnaire for the 2008-2012 MHSS clinical study is shown in Appendix C of the 2008-2012 MHSS operations report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2014a).

- WHODAS variables (IMPREMEM, IMPCONCN, IMPGOUT, IMPGOUTM, IMPPEOP, IMPPEOPM, IMPSOC, IMPSOCM, IMPHHLD, IMPHHLDM, MPRESP, IMPRESPM, IMPWORK).

These recoded item scores were used to calculate the total scores of K6 (K6SCMON, K6SCYR, K6SCMAX, WSPDSC2), WHODAS (WHODASC2, WHODASC3), and SDS (SDSSC2, SDSSC3). These total score variables are available for all adults on both the annual NSDUH public use file (CBHSQ, Substance Abuse and Mental Health Services Administration [SAMHSA], 2013) and the NSDUH 2008-2012 adult clinical interview data file (CBHSQ, SAMHSA, 2014b). See Appendices A, B, and C for more details on the K6, WHODAS, and SDS item score and total score variables.

Table 4.1 displays the unweighted numbers and percentages of the adult respondents in the 2008 to 2012 NSDUH main study who had at least one of the six past month K6 item scores missing, had at least one of the WHODAS or SDS item scores missing, and had their WHODAS or SDS questions skipped out because the sum of all K6 item scores was zero. These cases had their missing values set to zero.

An evaluation comparing the zero-assignment imputation method (i.e. assigning all the missing values with zeros) and weighted sequential hot deck imputation method resulted in no differences in estimates for SMI and any mental illness (AMI) (CBHSQ, in press b). Therefore, the recoded K6 and WHODAS total scores on the data file were used in the SMI prediction models.

Two other variables from the NSDUH main study—having past year major depression episode (AMDEYR) and having past year serious suicidal thoughts (MHSUITHK)—also had their missing values recoded as zeros: MHSUITHK was recoded as MHSUTK_U, and AMDEYR was recoded as AMDEY2_U. The recoded variables (MHSUTK_U and AMDEY2_U) were used in the 2012 revised models (see Section 7.3.4 in Chapter 7 for more details) as predictors for the mental illness status of NSDUH respondents. These recoded variables are also available for all adults on both the annual NSDUH public use file (CBHSQ, SAMHSA, 2013) and the NSDUH 2008-2012 adult clinical interview data file (CBHSQ, SAMHSA, 2014b). Table 4.2 reports the unweighted missing numbers and percentages of these two variables. The missing rates of both variables were all less than 1 percent for the 5 successive years. From 2008 through 2012, the prevalence rates of past year MDE were always less than 7 percent, and the rates of past year serious suicidal thoughts were always less than 4 percent. Therefore, assigning missing values with zeros will be very close to any other statistical imputation methods when estimating for SMI and AMI.

Table 4.1 Unweighted Missing Numbers and Percentages of K6, WHODAS, and SDS Item Scores: 2008-2012 All-Adult NSDUH Main Study

Variable	At Least One Item Was Missing		All Items Were Missing	
	Number	Percent (%)	Number	Percent (%)
K6 Item Scores				
2008-2012	2,498	1.09	456	0.20
2008A	237	1.05	45	0.20
2008B	249	1.08	46	0.20
2009	470	1.03	83	0.18
2010	525	1.15	98	0.21
2011	517	1.11	92	0.20
2012	500	1.09	92	0.20
WHODAS/SDS Item Scores, Skipped				
2008A-2012 (WHODAS)	-	-	37,296	18.06
2008A (WHODAS)	-	-	3,903	17.25
2008B (SDS)	-	-	3,922	17.02
2009 (WHODAS)	-	-	8,053	17.66
2010 (WHODAS)	-	-	8,313	18.13
2011 (WHODAS)	-	-	8,658	18.58
2012 (WHODAS)	-	-	8,369	18.26
WHODAS/SDS Item Scores, Non-Skipped				
2008A-2012 (WHODAS)	2,095	1.01	400	0.19
2008A (WHODAS)	221	0.98	50	0.22
2008B (SDS)	307	1.33	123	0.53
2009 (WHODAS)	482	1.06	79	0.17
2010 (WHODAS)	487	1.06	100	0.22
2011 (WHODAS)	472	1.01	84	0.18
2012 (WHODAS)	433	0.94	87	0.19

2008A = 2008 sample A; 2008B = 2008 sample B; - (i.e., a hyphen) = not applicable; K6 = Kessler-6, a 6-item psychological distress scale; NSDUH=National Survey on Drug Use and Health; SDS=Sheehan Disability Scale; WHODAS = World Health Organization Disability Assessment Scale.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 4.2 Unweighted Missing Numbers and Percentages of MDE and Suicidal Thoughts: 2008-2012 All-Adult NSDUH Main Study

Variable	Number	Percent (%)
Past Year MDE		
2008-2012	1,847	0.80
2008A	167	0.74
2008B	190	0.82
2009	353	0.77
2010	385	0.84
2011	377	0.81
2012	375	0.82
Past Year Suicidal Thoughts		
2008-2012	805	0.35
2008A	83	0.37
2008B	72	0.31
2009	147	0.32
2010	163	0.36
2011	162	0.35
2012	178	0.39

2008A = 2008 sample A; 2008B = 2008 sample B; MDE = major depressive episode; NSDUH=National Survey on Drug Use and Health.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

5. Sample Weighting

5.1 Background

The principal purpose of the subsample of clinical data collected as part of the Mental Health Surveillance Study (MHSS) was for use in estimating the prevalence of serious mental illness (SMI) among adults aged 18 or older. With the clinical data, a statistical model was developed in 2008 and revised in 2012 to predict whether an adult respondent to the main National Survey on Drug Use and Health (NSDUH) interview had SMI based on his or her main NSDUH interview responses. Sections 7.2 and 7.3 in Chapter 7 provide the details of the original 2008 SMI prediction model and the 2012 revised model. Although the weights for the clinical sample were created primarily for modeling SMI, another useful application was in providing nationally representative direct estimates of mental disorders.

The weights used in the 2008 modeling process were developed using fairly simple methods because of the relatively small sample size in a single year (for more details, see Chapter 4 in Liao et al., 2012). By contrast, with the availability of approximately 5,500 clinical interview cases accumulated from 2008 to 2012, more sophisticated weighting procedures could be developed. These procedures attempted to correct for two types of potential biases: (a) coverage biases resulting from not including in the clinical sample adults who responded to the main NSDUH interview in Spanish and (b) nonresponse biases resulting either from adults selected for the clinical follow-up study refusing to be recontacted or from those initially willing to be recontacted not providing useful clinical data (usually due to noncontact). Research has shown that the causes of these two types of nonresponse can be very different. The weighting adjustment procedures also incorporated techniques that increased the statistical efficiency (decreased the standard errors [SEs]) of direct estimates of mental disorders and were intended to improve the efficiency of SMI modeling as well.

This chapter describes how the final weights were created for use in the 2012 model revision process and for use in creating direct estimations based on the combined 5-year clinical sample from 2008 to 2012. The weights incorporated improved coverage, nonresponse, and poststratification adjustments that were performed separately for each year. The two 2008 clinical samples, one for respondents who were given World Health Organization Disability Assessment Schedule (WHODAS) questions in the NSDUH main interview (the 2008A sample) and one for respondents who were given Sheehan Disability Scale (SDS) questions (the 2008B sample), were treated as if they represented separate years (2008A and 2008B). The adjusted weights for each year were then combined using scaling factors for analyzing either the combined 2008-2012 or 2008A-2012 data.

Section 5.2 presents an overview of the revised weighting adjustments. Section 5.3 first provides comparisons on items from the main NSDUH interview between Hispanic respondents interviewed in English and Hispanic respondents interviewed in Spanish, then develops an

adjustment for this source of undercoverage.[20] Section 5.4 provides comparisons between respondents and nonrespondents at two different phases of nonresponse to the clinical interview. It then develops separate nonresponse adjustments for each phase. Section 5.5 first discusses the poststratification adjustment, then describes the scaling factors applied to the weights in order to increase the statistical efficiency when analyzing the combined 5-year data. Section 5.6 provides assessments of the overall impact of the adjustments, while Section 5.7 summarizes the new weighting process and its impact.

5.2 Components of the MHSS Analysis Weights

The weights for the 2008 modeling process were based on four adjustment factors: (1) the respondent's NSDUH adult-level analysis weight, (2) the inverse of the probability that the respondent was selected for the clinical interview, (3) a nonresponse adjustment, and (4) a poststratification adjustment. The NSDUH adult-level analysis weight, ANALWT, has 15 weight components. Each weight component represented either the selection probability at each selection stage or an adjustment on nonresponse, poststratification, or extreme weights. See Chen et al. (2014) for further details about the construction of NSDUH's adult-level analysis weight.

Out of necessity, weighting adjustments for the 2008 clinical data were based on a relatively small sample of roughly 1,500 clinical interviews (around 750 for each half sample). Subsequent years of data collection from 2008 to 2012 have resulted in the collection of about 5,500 clinical interviews overall. The larger sample size permitted a more detailed evaluation of the potential for bias due to undercoverage and nonresponse and the development of adjustments that can account for the specific undercoverage of Hispanics due to the clinical follow-up study only being conducted in English and the two phases of nonresponse in the clinical follow-up (i.e., at the end of the NSDUH main interview and during the attempt to administer the clinical interview).

The revised weighting adjustments incorporated three features designed to further reduce nonresponse bias and undercoverage. First, a separate coverage adjustment was added in order to focus on the undercoverage of Hispanics who completed the main NSDUH interview in Spanish, (see footnote 20; Section 3.8.2 provides more details on the noncoverage of non-English speakers).[21] Hispanic respondents who were eligible for the clinical follow-up had their weights adjusted to compensate for those who completed the interview in Spanish. This was done by forcing the total of the adjusted weights among Hispanics to equal the total of the main NSDUH Hispanic weights, both overall and within categories shown to be related to whether a Hispanic NSDUH respondent completed the main NSDUH interview in English. An example of such a category is adults with less than a high school education.

[20] The NSDUH main sampling weights are adjusted for undercoverage. The sole source of undercoverage in the clinical follow-up is the removal of adults responding to the main NSDUH interview in Spanish from the target population for clinical subsampling.

[21] By contrast, the coverage adjustment in the original weighting (i.e., the weighting used with the-2008 modeling process) occurred in the poststratification step. Although that step did correct for the general undercoverage of Hispanic adults, it did not try to account for the characteristics of those Hispanics who completed the main NSDUH interview in Spanish.

The second feature of the new weighting scheme employed two steps in nonresponse adjustment rather than a single step. The two steps corresponded with the two distinct phases of nonresponse in the clinical follow-up study: (1) initial nonresponse resulting from failing to agree to a clinical follow-up (see Section 5.4.1 for more details), and (2) final nonresponse from not completing the clinical follow-up after having agreed to do it earlier (see Section 5.4.1 for more details). Weights were first adjusted for initial nonresponse using variables found to be correlated with initial refusal at the end of the main NSDUH interview and expected to be related to key outcome measures of the Structured Clinical Interview for DSM-IV (SCID).[22] After this first step, the weights were adjusted again for final nonresponse using variables found to be correlated with final nonresponse and expected to be related to SCID measures.

The first three weighting adjustments were designed to remove biases due to undercoverage or nonresponse. The next weighting adjustment was a poststratification adjustment that was implemented to increase the statistical efficiency of estimators computed using the final adjusted weight. This adjustment forced the sum of the final weights for the clinical data to equal the sum of the final weights for the adult NSDUH main interview sample for a set of variables and variable interactions that predicted SMI.

Overall, the final revised weight (MHFNLWGT[23]) was the product of seven factors:

1. the respondent's NSDUH personal-level analysis weight (ANALWT);
2. a coverage adjustment to account for NSDUH respondents who completed that survey in Spanish;
3. the inverse of the probability that the respondent was selected for the clinical interview after completing the main NSDUH interview;
4. a nonresponse adjustment to account for those who, immediately after completing the main NSDUH interview, refused to be recontacted for a clinical interview;
5. a second nonresponse adjustment to account mostly for those who agreed to be recontacted but were unavailable for the clinical interview (this included a few who agreed to be recontacted for the clinical interview, but refused to respond when recontacted);
6. a poststratification or population weighting adjustment using control totals from the main NSDUH interview; and
7. a scaling factor used to combine clinical data across years efficiently.

Factors 2, 4, and 5 adjusted the weights to remove potential biases due to nonrandom exclusions and nonresponse. Factor 2 compensated for the exclusion of adults responding to the NSDUH main survey in Spanish from the target population for the clinical sample, while factors 4 and 5 compensated for different types of nonresponse to the clinical study. The weighting

[22] See Chapter 1 for details on the Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP or SCID) (First et al., 2002).

[23] This is the weight for the combined 2008-2012 clinical data. For the data excluding the 2008B sample, the final weight is MHFAAWGT.

adjustments were arranged in the same sequence as the corresponding rounds of exclusions, subsampling, and nonresponse. Beginning with the adult respondents to the NSDUH main survey (factor 1), adults responding in Spanish were first removed (causing the compensation adjustment in factor 2). This was followed by the probability selection of the clinical sample (captured in factor 3), the refusal of some sampled adults to be recontacted (factor 4), and so forth.

Table 5.1 provides a reference for the variable names given to various adjusted weights and adjustment factors used with the clinical data in the NSDUH 2008-2012 adult clinical interview data file and some mathematical denotations that are used in this report. The weighting adjustments in steps 4, 5, and 6 of the table were each composed of two adjustment factors, an initial trimming of a few large weights followed by a weighting adjustment that compensated for the trimming as it accomplished its designed task. Details on these steps are contained in Sections 5.4.3, 5.4.5, and 5.5.1, respectively, while step 2 is the focus on the next section.

5.3 Adjustment of Undercoverage of Hispanics

Because the clinical follow-up interview was conducted in English, adults who completed the main NSDUH interview in Spanish had no chance of being selected for the clinical interview. To the extent that differences exist on the main NSDUH interview measures between those who completed the main NSDUH interview in Spanish rather than in English, the fact that respondents in the clinical follow-up must have completed the NSDUH main interview in English could result in the estimates based on the clinical sample not being fully representative of the U.S. general population. The social stigma associated with mental health issues and alcohol and other substance use has been known to lead to underreporting of such sensitive behaviors among Hispanics, relative to non-Hispanics (e.g., Clark & Hill, 1991; Greenfield & Kaskutas, 1998; Maddahian, Newcomb, & Bentler, 1988; Silva de Crane & Spielberger, 1981). Additionally, the language of survey administration has been found to cue cultural values and norms among bilingual and bicultural respondents (Ross, Xun, & Wilson, 2002; Trafimow, Silverman, Mei-Tai Fan, & Shui Fun Law, 1997; Triandis, Davis, Vassiliou, & Nassiakou, 1965), suggesting that Hispanic respondents may answer culturally sensitive questions differently depending on whether they are asked in Spanish or in English.

Table 5.2 presents the number of Hispanic respondents in the adult NSDUH main study along with the numbers and percentages of those interviewed in Spanish by year of data collection. Although only around 4 percent of the NSDUH adult interviews between 2008 and 2011 were conducted in Spanish, over 25 percent of the adult Hispanic NSDUH respondents were interviewed in Spanish (roughly 35 percent when the numbers are weighted).[24]

[24] Because less than 1 percent of the adults who completed a main NSDUH interview in Spanish were not Hispanic, only the analysis weights of Hispanics were adjusted to compensate for NSDUH respondents who completed the survey in Spanish.

Table 5.1 Adjusted Weights and Adjustment Factors in the NSDUH 2008-2012 Adult Clinical Interview Data File

Step	Adjusted Weight	Initial Weight(s)	Adjustment/ Weighting Factor	Reason for Adjustment
1	ANALWT_A (w_k)	ANALWT	N/A	For 2009-2012, ANALWT_A = ANALWT; for 2008A and 2008B, ANALWT_A is used when each half sample is treated as a separate year.
2	ANALWT_E[1]	ANALWT_A	MHADJ_1 (A_{1k})	Compensates for NSDUH respondents who completed the NSDUH main interview in Spanish.
3	MHDSNWT[1] (F_k)	ANALWT_E[1]	MHWT1	Incorporates the clinical sample selection probabilities conditional on the eligible NSDUH respondents.
4	MHWTNR1[1]	MHDSNWT[1]	MHADJ_2a (A_{2ak}) MHADJ_2b (A_{2bk})	Compensates for selected individuals who do not agree to be recontacted for the follow-up clinical interview; MHADJ_2a is the trimming factor; MHADJ_2b is the adjustment factor.
5	MHWTNR2[1]	MHWTNR1[1]	MHADJ_3a (A_{3ak}) MHADJ_3b (A_{3bk})	Compensates for the remaining nonrespondents who initially agreed to be recontacted, but were unable to participate in the clinical interview due to other factors. MHADJ_3a is the trimming factor; MHADJ_3b is the adjustment factor.
6	MHNEWWGT	MHWTNR2[1]	MHADJ_4a (A_{4ak}) MHADJ_4b (A_{4bk})	Forces the sum of weights for the clinical data to match the sum of weights for the adult NSDUH main interview sample.
7	MHFNLWGT	MHNEWWGT	Scaling Factors for the 2008-2012 data	Rescales weights by year for getting SCID-based estimates of the 2008-2012 data.
7a	MHFAAWGT	MHNEWWGT	Scaling Factors for the 2008A-2012 data	Rescales weights by year for getting SCID-based estimates of the 2008A-2012 (WHODAS) data.

N/A = not applicable; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: MHFNLWGT = ANALWT_A * MHADJ_1 * MHWT1 * MHADJ_2a * MHADJ_2b * MHADJ_3a * MHADJ_3b * MHADJ_4a * MHADJ_4b * Scaling Factors for the 2008-2012 data.

[1] These variables are not available in the NSDUH 2008-2012 adult clinical interview data file, but they can be derived by multiplying ANALWT_A with related adjustment factors. For example, ANALWT_E = ANALWT_A * MHADJ_1.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table 5.2 Hispanic Respondents, by Language Version: 2008-2011 All-Adult NSDUH Main Study

Year/Sample	All Hispanic Respondents	Hispanic Respondents Interviewed in Spanish	
	Count	Count	Weighted Percent
2008A	3,369	1,024	34.9
2008B	3,392	1,055	36.8
2009	6,861	1,899	34.6
2010	6,889	1,883	34.5
2011	6,960	1,865	35.7
Combined	**27,471**	**7,726**	**35.2**

NSDUH = National Survey on Drug Use and Health.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2011.

5.3.1 Descriptive Characteristics of Hispanics Responding in English and Spanish

To better understand the consequences of excluding NSDUH respondents interviewed in Spanish for the clinical follow-up study, the possibility for coverage bias is examined by first comparing responses to the NSDUH mental health and substance use measures among Hispanics responding to NSDUH in English and those responding in Spanish. Differences in demographic characteristics and responses to mental health and substance use measures were examined between Hispanic respondents interviewed in Spanish and English from 2008 to 2011 (see Tables D1a to D1c in Appendix D). A small number (30) of non-Hispanic adult respondents were also interviewed in Spanish in the 2008 to 2011 NSDUHs. Those respondents were excluded from these language comparisons. The weights used for the comparisons were the NSDUH analysis weights (ANALWT_A).

Hispanic respondents interviewed in English were significantly different from Hispanic respondents interviewed in Spanish on a number of demographic characteristics (see Table D1a):

- Hispanics interviewed in Spanish tended to be older than those interviewed in English. Only 10.5 percent of those interviewed in Spanish were between the ages of 18 and 25, while 25.2 percent of those interviewed in English were similarly aged.
- Hispanics interviewed in Spanish were less educated than those interviewed in English. For example, 62.5 percent of those interviewed in Spanish reported having less than a high school education as compared with 21.3 percent of Hispanics interviewed in English.
- Hispanics interviewed in Spanish had significantly higher representation in the lowest income brackets. For example, 37.7 percent of those interviewed in Spanish reported an income of less than $20,000 as compared with 20.0 percent of those interviewed in English.
- Those interviewed in English were more likely to be born in the United States relative to those interviewed in Spanish (69.4 vs. 4.1 percent).
- A significantly higher percentage of Hispanics interviewed in English reported having health insurance coverage (71.1 percent) compared with those interviewed in Spanish (45.2 percent).

- A higher percentage of Hispanic respondents interviewed in English reported having no difficulty understanding the main NSDUH interview (88.3 percent) relative to those interviewed in Spanish (78.7 percent).
- Consistent with the literature on mental health issues, a significantly lower percentage of Hispanics interviewed in Spanish reported having been diagnosed with depression in their lifetime (5.3 percent) compared with Hispanics who were interviewed in English (8.9 percent).
- Spanish-language interviews with Hispanics yielded significantly lower estimates on mental health measures compared with English-language interviews with Hispanics, including Kessler-6 (K6) and WHODAS scores, lifetime and past year major depressive episode (MDE; i.e., depression), mental health treatment, suicidal thoughts, and treatment for attempted suicide (Table D1b). Specifically, the average K6 score among Hispanics responding in English was 5.1, while the mean for those responding in Spanish was 3.3. Similarly, past year MDE was 6.4 percent among Hispanics responding in English and 3.3 percent among those interviewed in Spanish.[25]
- Hispanics interviewed in English showed higher lifetime and past year prevalence rates than those interviewed in Spanish (Table D1c). Specifically, 50.1 percent of Hispanics responding in English reported any illicit lifetime drug use, while only 17.0 percent of Hispanics interviewed in Spanish reported such use.[26]

These results suggest that estimates related to mental health or substance use that are based on the clinical sample may be biased because of the exclusion of Hispanics who were interviewed in Spanish. The next section describes a method of adjusting for this source of undercoverage.

5.3.2 Development of the Coverage Adjustment to Compensate for Adults Interviewed in Spanish

The NSDUH weight (ANALWT) for an Hispanic adult respondent *k* to the clinical follow-up interview was adjusted yearly by a factor (MHADJ_1) of the form:

$$A_{1k} = \frac{1+\exp(\mathbf{g}^T\mathbf{z}_k)}{1+\exp(\mathbf{g}^T\mathbf{z}_k)/7}. \qquad (5.1)$$

The adjustment factor in equation (5.1) comes from the implicit estimation of the probability that an Hispanic NSDUH respondent was interviewed in English based on his or her characteristics. Those characteristics are put into mathematical form by the components of the

[25] Further language comparisons focused on NSDUH's mental health measures are available in Table D1b of Appendix D.

[26] Language comparisons focused on NSDUH's more specific substance use measures are available in Table D1c of Appendix D.

vector $\mathbf{z}_k$. The estimated probability for respondent k with the characteristics captured in $\mathbf{z}_k$ is $1/A_{1k}$. Equation (5.1) implies that this estimated probability cannot be less than 1/7.

Using combined NSDUH data from 2008 to 2011, a weighted (by ANALWT_A) logistic regression analysis among Hispanics was used to pare down the choice of candidates for $\mathbf{z}_k$ from a much larger potential list.[27] The dependent variable in this analysis was whether a Hispanic NSDUH respondent completed the survey in English. Predictor variables in the regression analysis were selected as components of $\mathbf{z}_k$ when they were significant at an alpha level of 0.02.

The components of $\mathbf{z}_k$ finally selected for equation (5.1) were

- a categorical age group variable (18 to 25, 26 to 34, and 35 or older);
- a categorical education-level variable (less than high school, high school graduate, other);
- a categorical variable for the number of years the respondent lived in the United States (born in the United States, else < 5 years, else < 10 years, else);
- an indicator of whether the respondent was alone when answering the main NSDUH interview; and
- the respondent's NSDUH K6 score.[28]

Using the WTADJUST procedure in SUDAAN 11.0 (RTI International, 2012),[29] a parameter vector $\mathbf{g}$ in equation (5.1) was found yearly that satisfied the calibration equation:

$$\sum_S w_k \mathbf{z}_k = \sum_S w_k A_{1k} \mathbf{z}_k, \tag{5.2}$$

where S was the yearly NSDUH adult respondent sample, and w_k was the NSDUH analysis weight for respondent k. When k completed the main NSDUH interview in Spanish, A_{1k} was defined to be 0.[30] This allowed both sides of the equality in equation (5.2) to be summed across the entire NSDUH respondent sample. Equations (5.1) and (5.2) forced the weighted eligible-adult NSDUH sample for a clinical interview in a year to look like the full adult NSDUH main interview sample with respect to the components of the vector $\mathbf{z}_k$. Note that a different $\mathbf{g}$ is estimated every year using the same vector of components.

---

[27] These candidates were gender, five age group levels, age (continuous), four education levels, years lived in the United States (four categories), privacy when answering NSDUH, WHODAS score (continuous), adjusted WHODAS score (continuous), K6 score (continuous), adjusted K6 score (continuous), NSDUH lifetime MDE (binary), NSDUH past year MDE (binary), and suicidal thoughts (binary). See Chapter 2 for a discussion of the alternative K6 and WHODAS scores. For the 2008B sample, WHODAS and alternative WHODAS scores were replaced by an SDS measure. Hispanic respondents interviewed in English were found to be significantly different from Hispanic respondents interviewed in Spanish on these candidate variables.

[28] K6SCMAX is the respondent's K6 score (0 to 24) for his or her worst month in the past year.

[29] The generalized exponential model (GEM) macro, which had been used in adjusting the main NSDUH's person-level weights (see Chen et al., 2014), could have been employed in place of WTADJUST to produce the same results.

[30] This treats undercoverage like nonresponse and sets the lower bound at 1 and the upper bound at 7.

The **g** that satisfies equation (5.2) is a consistent estimator for **γ** when the probability (i.e., propensity) that a Hispanic *k* responds in English has the following truncated logistic form:

$$\rho_k = \frac{1+\exp(\boldsymbol{\gamma}^T \mathbf{z}_k)/7}{1+\exp(\boldsymbol{\gamma}^T \mathbf{z}_k)}.$$

Were 7 replaced by infinity and γ by –γ in the above equation, ρk would have the standard logistic form. Using the calibration equation to find **g** is an alternative method to (weighted) maximum likelihood.

To show how the coverage adjustment modified the NSDUH adult-level weights so that the Hispanics who responded to the NSDUH main interview in Spanish (and were excluded from the clinical sample) were represented by Hispanics with otherwise similar characteristics who responded to the NSDUH main interview in English (and were included in the clinical sample), it is helpful to divide adult Hispanic respondents to the NSDUH main survey into five *propensity strata* based on each respondent's implicitly estimated propensity (i.e., probability) of responding in English.[31]

$$p_k = \frac{1+\exp(\mathbf{g}^T \mathbf{z}_k)/7}{1+\exp(\mathbf{g}^T \mathbf{z}_k)}.$$

The five propensity strata are the adults with the highest 60 percent of p_k values, next 20 percent, next 10 percent, next 5 percent, and lowest 5 percent. Table 5.3 displays the ratio of the weighted population totals for adults after the coverage adjustment to before the coverage adjustment in each propensity stratum. The ratio is close to 1 for the highest response propensity stratum with 60 percent (11,847 divided by 19,745) of the unweighted sample of Hispanics who responded to the NSDUH main interview in English, but it is over 6 for the lowest response propensity stratum. This means that the weights for those adults most likely to have answered the NSDUH main interview in Spanish increased roughly 600 percent on average because of the coverage adjustment, while they increased hardly at all for over 80 percent of the adult Hispanics. Compared with the higher propensity strata, those in the lowest propensity strata tended to have characteristics more similar to Hispanics who responded to the main NSDUH interview in Spanish (i.e., those excluded from the clinical interview). Hence, increasing the weights for the lowest propensity stratum compensates for the undercoverage of the excluded Hispanics.

5.4 Nonresponse Adjustment

5.4.1 Components of Nonresponse in the MHSS Clinical Follow-Up

Nonresponse to the MHSS clinical interview consists primarily of two components that reflect the two-stage process of participation in the survey. First, a respondent selected from the

[31] In principle, all adult Hispanics have a probability-of-responding-in-English of this form, although only those who actually responded to the NSDUH main interview in English were assigned here to propensity strata.

main NSDUH interview to participate in the MHSS clinical follow-up was asked whether or not he or she would like to participate in the clinical interview, as Exhibit 5.1 shows.

Table 5.3 Summary of Coverage Adjustments for Excluding Spanish-Language Version Respondents, by Propensity Strata: 2008 to 2011 NSDUH

Propensity Stratum	Sample Size	Weighted Population Estimate before Adjustment (in Thousands)	Weighted Population Estimate after Adjustment (in Thousands)	Ratio of after to before Weighted Population Estimates
Hispanics Responding in English, by Modeled Propensity to Respond in English	19,745	88,197	127,311	1.4435
Very High (Upper 60%)	11,847	39,114	39,150	1.0009
High (Next 20%)	3,943	19,531	20,089	1.0286
Moderate (Next 10%)	1,979	13,005	17,607	1.3539
Low (Next 5%)	988	5,317	16,513	3.1057
Very Low (Lowest 5%)	988	5,569	33,953	6.0972
Hispanics Responding in Spanish	7,726	44,776	0	0.0000
All Hispanic Respondents	**27,471**	**127,311**	**127,311**	**1.0000**

NSDUH = National Survey on Drug Use and Health.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2011.

Exhibit 5.1 Recruitment Screen for 2008-2012 MHSS Clinical Follow-Up

RECRUIT1

[IF ELIGIBLE FOR MENTAL HEALTH FOLLOW-UP STUDY] You have been randomly selected to participate in one additional study for the U. S. Public Health Service. This interview will ask questions about mental health issues. It will be conducted over the telephone and will take about an hour. Participation in this interview is voluntary and all of your answers will be confidential.

If you agree to participate, I will pay you an additional $30 today. Within the next two weeks, a different interviewer will call you to explain more about the interview and to schedule a convenient time to complete it. If you wish, you may complete the full interview when the interviewer calls.

HAND FOLLOW-UP STUDY DESCRIPTION TO RESPONDENT. Please read this statement. It describes the survey and the legislation that assures the confidentiality of any information you provide.

1 RESPONDENT AGREES TO RECONTACT
2 RESPONDENT DOES NOT AGREE TO RECONTACT
3 RESPONDENT IS NOT AVAILABLE DURING THE SPECIFIED TIME PERIOD.

RECRT4WK

[IF RECRUIT1=3] To accommodate your schedule, an interviewer will be available to call you about this study and schedule a convenient time to complete the interview within the next four weeks.

1 RESPONDENT AGREES TO RECONTACT
2 RESPONDENT DOES NOT AGREE TO RECONTACT
3 RESPONDENT IS NOT AVAILABLE DURING THE SPECIFIED TIME PERIOD.

Second, nonresponse occurs when the clinical interviewer attempts to contact the respondent who agreed to participate at the end of the main NSDUH interview. If there are very few differences between the nonrespondents in these two phases of participation in the MHSS clinical interview, they can be treated as a homogenous group in nonresponse adjustments. However, if there are differences between these groups of nonrespondents, adjustments tailored to each phase of nonresponse may have greater potential for reducing nonresponse bias.

Further distinctions among nonrespondents in the MHSS clinical study are possible but were not considered for evaluating differences between respondents and nonrespondents or for use in nonresponse weighting adjustments due to small sample sizes. Table 5.4 shows the distribution of result codes for nonrespondents for the clinical follow-up for 2008 to 2011. These are adults selected for the MHSS clinical follow-up who agreed to participate when asked at the end of the main NSDUH interview but then did not complete the clinical interview.[32]

Table 5.4 Result Codes for Adults Selected for Clinical Follow-Up Interviews and Agreeing to Participate, 2008 to 2011

Result Code	Number of Adults Selected	Percentage
Code 73 (Breakoff, Partial Interview)	157	14.2
Code 74 (Unable to Contact)	743	67.5
Code 75 (Phone Number Problem)	129	11.7
Code 76 (Refusal)	27	2.5
Code 77 (Other)	44	4.0
Total Incompletes	**1,100**	**99.9**

NSDUH = National Survey on Drug Use and Health.

NOTE: Due to rounding, the percentages do not add to 100 percent.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2011.

Most of the nonresponse at the point of attempting to conduct the clinical interview (67.5 percent) was classified as unable to contact. However, this classification does not mean that in all of these cases contact was not made with the selected adult. Many of these may have been passive refusals in which respondents did not actually refuse to participate but may have simply avoided subsequent contacts with the clinical interviewer or were not able to complete the interview within the 4 weeks allowed for completing clinical interviews. More important, only 2.5 percent of the nonresponse at this stage consisted of refusals. Thus, the causes of nonresponse at this stage can be different from that at the previous stage at the end of the main NSDUH interview, in which nonrespondents immediately refused to participate.

[32] Table 5.4 also includes clinical interview cases where the record was determined to be falsified (discussed in Section 3.8.4 in Chapter 3) and cases in 2008 through 2011 that should have been selected for the clinical interview based on their characteristics but were assigned a zero probability of selection instead. For more details, see Section 3.8.1 in Chapter 3.

5.4.2 Descriptive Characteristics of Adults Agreeing and Not Agreeing to Participate in the Clinical Follow-Up

Tables D2a to D2c in Appendix D present comparisons between NSDUH respondents agreeing to take part in the clinical follow-up and those who refused to do so. The weight used for these comparisons was the full base weight MHDSNWT for the clinical sample, which was the weight in the previous step that accounted for undercoverage of Hispanics interviewed in Spanish (ANALWT_E), adjusted for clinical sample selection probabilities.

Descriptive statistics comparing adults who agreed to participate in the clinical interview with initial nonrespondents are available from Table D2a:

- Those who agreed to participate in the clinical interview were generally younger than those who declined to participate. For example, 18.4 percent of those agreeing to participate were between the ages of 18 and 25, while only 8.6 percent of those refusing to participate were in that age range. Also, 51.6 percent of those refusing to participate were aged 50 or older, while only 38.3 percent of those who agreed to participate were aged 50 or older.
- Blacks made up 10.2 percent of those agreeing to participate, while blacks only made up 4.6 percent of those who refused to participate.
- An estimated 26.1 percent of adults who agreed to participate reported some college as their highest level of education as compared with 17.9 percent of adults refusing to participate in the clinical follow-up. There were no other statistically significant differences in the other levels of educational attainment.
- In terms of employment status, those refusing to participate were much more likely to be retired than those agreeing to participate (22.3 vs. 10.7 percent). Also, adults agreeing to participate were more likely to not be currently working but to have worked in the past 12 months than those refusing to participate (8.1 vs. 3.8 percent).
- Those refusing to participate were less likely to report ever having been diagnosed with depression (5.9 percent) than those agreeing to participate (15.1 percent). Also, those refusing to participate were more likely to report ever having been diagnosed with ulcers than those agreeing to participate (5.6 vs. 0.9 percent).
- Adults refusing to participate had shorter interview times than those agreeing to participate. For example, 49.9 percent of those refusing to participate completed the main NSDUH interview in less than 60 minutes as compared with only 27.7 percent of those who agreed to participate. Also, 77.2 percent of those agreeing to participate completed the adult depression module in less than 2 minutes, while 92.5 percent of those refusing to participate completed the depression module in less than 2 minutes.

Mental health measures are presented in Table D2b for comparisons of adults who agreed to participate in the clinical interview with initial nonrespondents:

- Consistent with the differences in the time needed to complete the adult depression module, 7.4 percent of those agreeing to participate reported MDE in the past year, while only 3.1 percent of refusers reported past year MDE.

- Also, 15.2 percent of those agreeing to participate reported receiving any mental health treatment in the past year, while only 8.7 percent of those refusing to participate reported receiving such treatment.

Substance use measures are presented in Table D2c for comparisons of adults who agreed to participate in the clinical interview with initial nonrespondents:

- Those refusing to participate had lower rates of lifetime and past year illicit drug use compared with those who agreed to participate in the clinical follow-up.
- For example, among the lifetime measures of illicit drug use, all differences between adults who refused to participate and those agreeing to participate were statistically significant, except for the nonmedical use of tranquilizers.

5.4.3 Development of the Adjustment for Adults Not Agreeing to Participate in the Clinical Follow-Up

Before adjusting for nonresponse, a few large values in the (coverage-adjusted) full base weights[33] (MHDSNWT) were trimmed so that (ideally) no single respondent in the clinical sample had more than a 4 percent impact on the direct estimates based on the clinical sample in a year and thereby undermined an assumption of large-sample probability-sampling theory.

A full base weight F_k greater than 4 percent of the sum of NSDUH analysis weights in the year was trimmed back to that value (T). This meant that the full base weight was multiplied by a trimming factor (MHADJ_2a):

$$A_{2ak} = min(1, T/F_k),$$

where A_{2ak} was defined to be 0 for NSDUH respondents not selected for the clinical interview.

The larger the yearly sample sizes and the less oversampling, the fewer the weights needing trimming. In total, nine records had their full base weights trimmed. Five were from 2009, two from 2008B, and one each from 2008A and 2010. The relatively large samples in 2011 and 2012 had none.

Selection for the follow-up clinical interview was done at the same time as the main NSDUH interview. A fraction of those selected for the clinical interview did not agree to be recontacted by telephone for a clinical interview. The initial refusal adjustment factor (MHADJ_2b) had a form very similar to equation (5.1):

$$A_{2bk} = \frac{1+\exp(\mathbf{g}^T\mathbf{z}_k)}{1+\exp(\mathbf{g}^T\mathbf{z}_k)/U}. \qquad (5.3)$$

[33] These base weights are the weights of the full clinical sample before nonresponse adjustment with the coverage adjustment for the clinical sample as well as all the weighting adjustments in the NSDUH main survey person-level analysis weights.

The U in equation (5.3) capped the size of the adjustment and thus the impact a single record could have on estimation. It was set at 5 for most years and at 5.5 for 2009. That is as low as U could be set for the calibration equation (5.4), given in the following text, to be satisfied.

The components of $\mathbf{z}_k$ were selected from a much larger set of candidate variables. For a variable from the main NSDUH interview to be a candidate, there had to be a significant difference between its mean computed in the clinical sample selected to participate in the clinical follow-up using combined 2008 through 2011 clinical data and its mean in the subset of adults who agreed to be recontacted.

Starting with these candidate variables and using the same combined 2008-2011 clinical data, backward selection was applied to a logistic-regression analysis predicting whether a NSDUH respondent selected for the clinical interview would agree to be recontacted. The variables ultimately selected to be components of $\mathbf{z}_k$ in equation (5.3) had p values of less than 0.1. These components included the K6 score (0 to 24). The other components were (0/1) indicators for the following characteristics:

- age less than or equal to 25 years,
- black,
- at least some college,
- retired,
- worked in the past 12 months,
- lifetime depression,
- lifetime ulcer or ulcers,
- NSDUH main interview length of less than 60 minutes,
- NSDUH adult depression module of less than 2 minutes,
- past year MDE,
- lifetime cocaine or crack or heroin use,
- lifetime lysergic acid diethylamide (LSD) use,
- lifetime Ecstasy use,
- lifetime nonmedical use of psychotherapeutics,
- lifetime nonmedical use of pain relievers,
- lifetime methamphetamine use, and
- received any mental health treatment in the past year.

The vector $\mathbf{g}$ was chosen using the WTADJUST procedure to satisfy the calibration equation:

$$\sum_S F_k \mathbf{z}_k = \sum_S F_k A_{2ak} A_{2bk} \mathbf{z}_k, \tag{5.4}$$

where S was again a yearly NSDUH adult respondent sample and, by definition, $A_{2bk} = 0$ for a NSDUH respondent who was either not selected for the clinical interview or who refused to be recontacted when selected. This vector consistently estimates $\boldsymbol{\gamma}$ in the following model for the probability that an adult in the clinical sample agrees to be recontacted:

$$\rho_k = \frac{1+\exp(\boldsymbol{\gamma}^T \mathbf{z}_k)/U}{1+\exp(\boldsymbol{\gamma}^T \mathbf{z}_k)},$$

where an adult k selected for the clinical follow-up sample agreed to be recontacted is $p_k = 1/A_{2bk}$ when $A_{2bk} > 0$.

Table 5.5 shows how the weighting to adjust for those adults who refused to be recontacted after being selected for the clinical sample related to the estimated probabilities of agreeing to be recontacted. The adults who agreed to be recontacted were divided into five propensity strata based on their estimated probabilities of agreeing to be recontacted: highest 25 percent, next 25 percent, next 25 percent, next 15 percent, and lowest 10 percent. The table displays the ratio of the weighted population totals for adults after the weighting adjustment to before the weighting adjustment in each of the propensity stratum. The ratio is close to 1 for 75 percent of the selected sample (i.e., for the three highest propensity strata, divide the sum of 1,317, 1,318, and 1,317 by 5,271) and is as high as about 2 for the lowest response propensity stratum. This means that the weights for those adults in the lowest propensity stratum increased roughly 200 percent on average due to the weighting adjustment, while they hardly increased at

Table 5.5 Summary of Nonresponse Adjustments for Agreement to Participate in the Clinical Follow-Up, by Propensity Strata: 2008 to 2011 MHSS

Propensity Stratum	Sample Size	Weighted Population Estimate before Adjustment (in Thousands)	Weighted Population Estimate after Adjustment (in Thousands)	Ratio of after to before Weighted Population Estimates
Adults Agreeing, by Propensity to Agree to Participate in Clinical Follow-Up	**5,271**	**720,273**	**930,942**	**1.2925**
Very High (Upper 25%)	1,317	115,794	115,906	1.0010
High (Next 25%)	1,318	104,325	107,478	1.0302
Moderate (Next 25%)	1,317	182,829	200,206	1.0950
Low (Next 15%)	791	173,321	222,686	1.2848
Very Low (Lowest 10%)	528	144,003	284,666	1.9768
Sample Adults Not Agreeing to Participate in Clinical Follow-Up	894	210,669	0	0.0000
Total Sample	**6,165**	**930,942**	**930,942**	**1.0000**

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2011.

all for 75 percent of the adults in the three highest propensity strata. Those in the lowest propensity stratum (least likely to agree to participate in the clinical interview) had characteristics that were more similar to those who refused to participate than those in the higher propensity strata. Thus, adjusting their weights compensates for those who refused to participate in the clinical interview.

5.4.4 Descriptive Characteristics of Adults Completing and Not Completing the Clinical Follow-Up

Tables D3a to D3c in Appendix D present comparisons between respondents who completed the clinical follow-up and those who had agreed to participate in the clinical follow-up but did not. The weight used for these comparisons was MHWTNR1, which was the weight in the previous step that accounted for nonresponse by those who refused to take part in the clinical follow-up at the conclusion of the main NSDUH interview. Below are some highlights of these comparisons, particularly with respect to measures used in the adjustment for those not completing the clinical follow-up and items in the 2012 revised model for predicting SMI.

Descriptive statistics comparing adults who initially agreed to participate in the clinical interview with final nonrespondents are available from Table D3a:

- Women were more likely to complete the clinical follow-up interview than men. For example, 53.6 percent of those who completed the clinical interview were women, while men made up 60.6 percent of those who did not complete the clinical interview.
- Those who did not complete the interview were younger than those who did. Adults between the ages of 18 and 30 years old made up 23.0 percent of those who completed the clinical interview, but 36.6 percent of those who did not complete the clinical interview.
- Those who did not complete the clinical follow-up interview were less likely to be white and more likely to be Hispanic than those who completed it. White respondents made up 73.3 percent of those who completed the interview, but only 55.2 percent of those who did not complete it. Also, 12.4 percent of those who completed the interview were Hispanic, while Hispanics made up 27.8 percent of those who did not complete it.
- Respondents to the clinical follow-up interview reported higher levels of education than nonrespondents. For example, 29.2 percent of those who completed the interview were college graduates, while 16.1 percent of those who did not complete it were college graduates. Similarly, only 9.4 percent of those who completed the interview had less than a high school education, while 31.1 percent of those who did not complete it had less than a high school education.
- An estimated 30.4 percent of those who did not complete the clinical follow-up interview were not covered by health insurance. In contrast, only 14.0 percent of those who did complete it did not have health insurance.

- Respondents to the clinical follow-up interview were more likely to report no difficulties in understanding the main NSDUH interview (93.7 percent) compared with those who did not complete it (82.0 percent).

Mental health measures are presented in Table D3b for comparisons of adults who agreed to participate in the clinical interview with final nonrespondents:

- This table contains many of the items used in the 2012 revised model for predicting SMI. It is noteworthy that for measures such as K6 scores, WHODAS scores, past year MDE, and serious thoughts about suicide in the past year, there were no statistically significant differences between those who completed the clinical interview and those who did not.

Substance use measures are presented in Table D3c for comparisons of adults who agreed to participate in the clinical interview with final nonrespondents:

- Those refusing to participate had lower rates of lifetime and past year illicit drug use compared with those who agreed to participate in the clinical follow-up interview.
- For example, among the lifetime measures of illicit drug use, all of the differences between adults who refused to participate and those who agreed to participate were statistically significant except for the nonmedical use of tranquilizers.

Thus far, differences between respondents and nonrespondents have been presented in this section and the previous one. Differences between the two types of nonrespondents also are worth noting. Initial nonrespondents (i.e., those who refused to participate in the clinical follow-up interview immediately at the end of the main NSDUH interview) had lower prevalences for a number of items related to mental health compared with nonrespondents who initially agreed to participate in the clinical follow-up interview but ultimately did not complete that interview (i.e., the final nonrespondents). For example, 5.9 percent of initial nonrespondents reported ever having been diagnosed with depression (Table D2a), while 13.0 percent of final nonrespondents did so (Table D3a). The mean K6 score for initial nonrespondents was 3.2 (Table D2b), while it was 4.9 for final nonrespondents (Table D3b). Past year and lifetime rates of MDE were lower for initial nonrespondents than for final nonrespondents (Tables D2b and D3b).

There were also demographic differences between the two types of nonrespondents. Initial nonrespondents were more likely to be female than final nonrespondents (52.1 vs. 39.4 percent; Tables D2a and D3a). Initial nonrespondents were older than final nonrespondents. Only 13.5 percent of initial nonrespondents were aged 18 to 30 years old, while 36.6 percent of final nonrespondents were in that age range. Initial nonrespondents reported higher levels of education than final nonrespondents; 27.5 percent of initial nonrespondents were college graduates, while 16.1 percent of final nonrespondents were college graduates.

5.4.5 Development of the Adjustment for Adults Not Completing the Clinical Follow-Up

The adjustments for nonresponse among adults who had originally agreed to be recontacted for clinical follow-up paralleled the refusal adjustments in the previous

adjustment step. First, seven records had their adjusted weights for initial refusal (MHWTNR1) trimmed to *T* (the trimming factor is MHADJ_3a). Three were from 2009, two from 2008B, and one each from 2008A and 2010.

The nonresponse adjustment factor (MHADJ_3b) for not completing the clinical interview employed an unbounded (by *U*) version of equation (5.1):

$$A_{3bk} = 1 + \exp(\mathbf{g}^T \mathbf{z}_k). \tag{5.5}$$

The components of $\mathbf{z}_k$ were selected for equation (5.5) from a large set of candidate variables (which were again chosen based on univariate comparisons) by fitting a logistic regression to adults who had agreed to be recontacted for the clinical interview from 2008 through 2011. The dependent variable for this model was whether or not a selected adult completed the clinical interview. Each component had a *p* value of less than 0.1.

The components of $\mathbf{z}_k$ were 0/1 indicators of the following characteristics:

- female,
- white,
- less than high school,
- covered by any health insurance,
- no difficulty understanding the main NSDUH interview,
- analgesic dependence or abuse,
- alcohol dependence or abuse,
- lifetime sedative use,
- past year cigarette use,
- NSDUH main interview of less than 60 minutes, and
- lifetime nonmedical use of psychotherapeutics.

The vector **g** was chosen using WTADJUST to satisfy the calibration equation:

$$\sum_S F_k A_{2ak} A_{2bk} \mathbf{z}_k = \sum_S F_k A_{2ak} A_{2bk} A_{3ak} A_{3bk} \mathbf{z}_k, \tag{5.6}$$

where *S* was again a yearly NSDUH adult respondent sample, A_{3ak} the trimming factor applied to the weight after initial refusal, if necessary, and $A_{3bk} = 0$ for any NSDUH respondent who was not a respondent in the clinical follow-up study. This vector consistently estimates $\boldsymbol{\gamma}$ in the following model for an adult in the clinical sample who agreed to be recontacted actually completing the clinical interview:

$$\rho_k = \frac{1}{1 + \exp(\boldsymbol{\gamma}^T \mathbf{z}_k)}.$$

The estimated probability that an adult k who agreed to be reinterviewed actually completed the clinical follow-up interview is $p_k = 1/A_{3bk}$ when $A_{3bk} > 0$.

Table 5.6 shows how the weighting to adjust for adults who agreed to be recontacted but failed to complete the clinical follow-up reinterview related to the probabilities of completing the clinical interview among those who agreed to be recontacted. The adults who completed the clinical interview were divided into five propensity strata based on their estimated probabilities of completing the clinical interview: highest 25 percent, next 25 percent, next 25 percent, next 15 percent, and lowest 10 percent. The table displays the ratio of the weighted population totals for adults after the weighting adjustment to before the weighting adjustment in each of the propensity strata. The ratio is close to 1 for 75 percent of the adults who completed the clinical follow-up interview (i.e., for the highest three propensity strata). By contrast, for the lowest propensity stratum, it is close to 3. This means that the weights for those adults in the lowest response propensity stratum increased roughly 300 percent on average because of the weighting adjustment, while they increased hardly at all for 75 percent adults in the highest three propensity strata. Adults in the lowest propensity stratum (i.e., those least likely to complete the clinical follow-up given that they initially agreed to participate) had characteristics most similar to adults who initially agreed but failed to complete the clinical follow-up. Thus, increasing the weights of adults in the lowest propensity stratum compensates for those clinical-interview nonrespondents.

Table 5.6 Summary of Nonresponse Adjustments for Adults Completing the Clinical Follow-Up Given That They Agreed to be Recontacted, by Propensity Strata: 2008 to 2011 MHSS

Propensity Stratum	Sample Size	Weighted Population Estimate Before Adjustment (in Thousands)	Weighted Population Estimate After Adjustment (in Thousands)	Ratio of After to Before Weighted Population Estimates
Adults Completing Clinical Follow-Up, by Propensity to Complete Clinical Follow-Up	**4,031**	**715,984**	**930,942**	**1.3002**
Very High (Upper 25%)	980	164,197	168,978	1.0291
High (Next 25%)	1,034	193,532	215,857	1.1154
Moderate (Next 25%)	1,009	206,169	252,779	1.2261
Low (Next 15%)	604	107,795	163,737	1.5190
Very Low (Lowest 10%)	404	44,291	129,591	2.9259
Sample Adults Initially Agreeing but Not Completing Clinical Follow-Up	1,240	214,958	0	0.0000
Total Sample	**5,271**	**930,942**	**930,942**	**1.0000**

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2011.

5.4.6 Summary of Nonresponse Adjustments

It is worth noting that there is very little overlap between the two sets of measures used to adjust for the two different phases of nonresponse. The measures that are predictive of agreement to participate in the clinical interview at the end of the main NSDUH interview are mostly different from the items predictive of completing the clinical follow-up interview, conditional on having agreed to participate. The only items in common between the two sets of variables used for nonresponse adjustments are race/ethnicity, education, NSDUH main interview length, and lifetime nonmedical use of psychotherapeutics. The adjustment for adults refusing to participate at the end of the main NSDUH interview contains 13 other variables, and the adjustment for not completing the clinical interview contains 7 other variables. This suggests that the causes of these two types of nonresponse are distinct from each other and that the adjustments should be carried out separately.

5.5 Development of Revised Poststratification Adjustment and Scaling

5.5.1 Poststratification

The term "poststratification" has been used to describe a weighting step that forces the weighted totals of a set of variables to equal the population totals provided from an outside source such as the U.S. Census Bureau. [34] For the NSDUH main survey and previously for the clinical sample under the 2008 MHSS modeling process (described in Liao et al., 2012, Chapter 4), such a poststratification step was used to adjust for undercoverage in the sampling frame. In developing the new clinical sample weights, however, it was assumed that both coverage errors and nonresponse biases were removed prior to conducting the poststratification step. The goal of this step was to decrease the SEs of estimates for mental health characteristics estimated directly from the clinical sample. This was done by forcing weighted totals after adjustment to equal totals not provided from an outside source per se but estimated from the overall NSDUH main survey adult respondents.

The poststratification process began with the adjusted weight for a completed clinical interview (MHWTNR2), $q_k = F_k A_{2ak} A_{2bk} A_{3ak} A_{3bk}$. First, six records had their weights trimmed to T as in the previous two steps. Three were from 2009 and one each from 2008A, 2008B, and 2010.

In previous adjustment steps, the adjustment factor implicitly estimated the probability of responding to the NSDUH main interview in English or the probability of responding to the clinical follow-up interview. In this poststratification step, no such probability is being estimated. In fact, the adjustment factor can be viewed as an "estimate" of 1. The poststratification weighting adjustment had the following form:

[34] See Chen et al. (2014) for a description of the poststratification step in developing the NSDUH main survey person-level analysis weights. This use of the term "poststratification" does not strictly conform to the textbook definition in, for example, Lohr (1999, pp. 113-114), which requires target population totals to be for mutually exclusive groups (and so limits the number of types of groups that could be used in the step).

$$A_{4bk} = \frac{\exp\left\{[(U_k - 1)/U_k]a_k \mathbf{g}^T \mathbf{z}_k\right\}}{(U_k - 1)/U_k + \exp\left\{[(U_k - 1)/U_k]a_k \mathbf{g}^T \mathbf{z}_k\right\}/U_k}, \quad (5.7)$$

where $a_k = q_k/w_k$ and $U_k = T/q_k$ in all years but 2008A, where it was relaxed to 1.25 T/q_k for the calibration equation (below) to hold.[35] This setting bounded the fully adjusted weights themselves (MHNEWWGT) to T where possible and 1.25 T otherwise. Unlike in the previous adjustments, the choice of A_{4bk} was not related to the probability that adult k responded to the NSDUH main survey in English or successfully completed the clinical sample. As a result, the bounding parameters U_k were allowed to vary from one adult to another.

In equation (5.7), the vector $\mathbf{z}_k$ consists of the variables chosen because their inclusion in $\mathbf{z}_k$ decreased the SEs of estimates for past year SMI, any mental illness (AMI), and MDE computed from the clinical sample without the aid of a model. These variables are as follows:

- indicators for six categories of gender (male and female) by age (18 to 25, 26 to 34, 35 or older) categories,
- indicators for race/ethnicity categories (Hispanic, non-Hispanic white, non-Hispanic black, other),
- an indicator for suicidal thoughts,
- indicators from the NSDUH main interview for a past year and lifetime MDE,
- the interaction term between an alternative K6 score (defined as the maximum value between 0 and K6SCMAX – 7)[36] and three age categories, and
- the interaction term between an alternative WHODAS score (ranges from 0 to 8) and three age categories.[37]

The vector $\mathbf{g}$ in equation (5.7) was found using the WTADJX[38] procedure in SUDAAN 11.0 (RTI International, 2012) to satisfy the calibration equation:

[35] Equation (5.7) follows the bounded pseudo-optimal calibration weighting discussed in Kott (2011).

[36] See Chapter 2 for a discussion of the alternative K6 score.

[37] See Chapter 2 for a discussion of the alternative WHODAS score. For the 2008B sample, it was replaced by an alternative SDS measure.

[38] The insertion of the a_k into equation (5.7), suggested in Kott (2011), tends to decrease the SEs of the resulting estimators (they would still decrease without the a_k, but not by as much). Because of it, WTADJUST could no longer be used (nor could the GEM macro) and was replaced by WTADJX. Implementing WTADJX with ADJUST=POST in this context requires some work. First, a dataset is created with both main NSDUH and clinical sample respondents. Adults in the latter group are listed twice in this dataset, once as members of the main sample and once as members of the clinical sample. A main sample member has an initial weight of w_k, while a clinical sample member has an initial weight of $\eta_{kk} = F_k A_{2ak} A_{2bk} A_{3ak} A_{3bk} A_{4ak}$. The components of $c_k\mathbf{z}_k$ are in the CALVARS statement, where $c_k = 1$ for the clinical sample members, and $c_k = -1$ for the main sample members. The MODEL statement is _ONE_ = the components of $m_k\mathbf{z}_k$, where $m_k = a_k$ for the clinical sample members, and $m_k = 0$ for the main sample members. The components of POSTWGT are all 0. The lower bound is 0, and the upper bound is U_k.

$$\sum_S w_k \mathbf{z}_k = \sum_S F_k A_{2ak} A_{2bk} A_{3ak} A_{3bk} A_{4ak} A_{4bk} \mathbf{z}_k, \quad (5.8)$$

where A_{4ak} was the trimming factor applied to the weight after other nonresponse adjustments, if necessary, and $A_{4bk} = 0$ for any NSDUH respondent who was not a respondent in the clinical follow-up study.

Table 5.7 presents the poststratification controls used in the adjustment process along with the ratio of the summed weights before and after adjustment. The ratio can be viewed as the relative degree to which weights were increased or decreased from the previous weighting step to meet the desired control totals as specified in the poststratification. The groups shown are those used in the poststratification. For targets that involved noncategorical measures (e.g., the K6, WHODAS, or SDS scores), means are provided with the age categories that were used in the adjustment as well as the ratios between these means before and after adjustment. For these, a ratio greater than 1 indicates that adults with higher K6, WHODAS, or SDS scores had their weights increased by the poststratification, while a ratio of less than 1 indicates that adults with higher scores on these measures had their weights decreased. For most groups, the ratios of the weighted population totals from after to before poststratification are fairly close to 1, indicating that on average, the cases in these groups did not have their weights increased or decreased by much due to the poststratification.

5.5.2 Scaling

Because of the variations in the clinical interview sample sizes, as well as the sample allocations and weighting adjustments, a scaling factor was applied to weights in each year before combining the years. This was applied to reduce the variance of the estimates for mental health characteristics estimated directly from the clinical diagnoses and to increase the efficiency in the modeling described in Chapter 7. The scaling factors were determined based on the reduction of the variance in the estimates of SMI, AMI, and past year MDE that were made directly from SCID diagnoses. For the combined 2008-2012 clinical sample including both the 2008A and 2008B clinical samples, analysis weights, designated as MHFNLWGT, were created by scaling the MHNEWWGT weights in each year with the following scaling factors: 6 percent for 2008A sample, 6 percent for 2008B sample, 4 percent for 2009, 14 percent for 2010, 35 percent for 2011, and 35 percent for 2012.

For the combined 2008-2012 clinical sample including only the 2008A sample, the weight MHNEWWGT was scaled by using the following scaling factors: 12 percent for the 2008A sample, 4 percent for 2009, 14 percent for 2010, 35 percent for 2011, and 35 percent for 2012. This set of scaled weights, designated as MHFAAWGT, was applied in the development of the 2012 model.

Table 5.7 Poststratification Controls and Average Adjustment: 2008 to 2011 MHSS Clinical Sample

Description	Value(s)	Weighted Population Estimate before Adjustment (in Thousands)	Weighted Population Estimate after Adjustment (in Thousands)	Ratio of after to before Weighted Population Estimates
Gender	Male	462,212	441,105	0.9543
	Female	468,731	472,923	1.0089
Age	18 to 25	163,236	134,892	0.8264
	26 to 49	381,976	395,813	1.0362
	50 or Older	385,731	383,322	0.9938
Gender, by Age	Male, 18 to 25	89,423	67,916	0.7595
	Male, 26 to 49	200,597	195,318	0.9737
	Male, 50 or Older	172,191	177,871	1.0330
	Female, 18 to 25	73,813	66,977	0.9074
	Female, 26 to 49	181,379	200,494	1.1054
	Female, 50 or Older	213,539	205,451	0.9621
Race	White	643,530	621,345	0.9655
	Black	99,531	105,144	1.0564
	Hispanic	126,260	127,311	1.0083
	Asian and Others	61,622	60,227	0.9774
Seriously Think about Killing Self in Past 12 Months	Yes	36,226	34,079	0.9407
Lifetime MDE	Yes	125,099	117,823	0.9418
Past Year MDE	Yes	65,985	60,811	0.9216
Overall		930,942	914,027	0.9818
Alternative K6 Score, by Age[1]	18 to 25	2.20	2.35	1.0711
	26 to 49	1.77	1.62	0.9167
	50 or Older	0.94	0.88	0.9358
Alternative WHODAS Score, by Age[2]	18 to 25	1.27	1.18	0.9311
	26 to 49	1.04	0.93	0.8958
	50 or Older	0.55	0.66	1.2038
Alternative SDS Score, by Age[3]	18 to 25	0.29	0.40	1.3753
	26 to 49	0.26	0.34	1.2807
	50 or Older	0.24	0.19	0.7938

K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SDS = Sheehan Disability Scale; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: The weighted population estimates before adjustment used WEIGHT = MHWTNR2; the weighted population estimates after adjustment used WEIGHT = MHNEWWT.

[1] Alternative K6 score is defined as max(0, K6SCMAX – 7).
[2] Alternative WHODAS total score (range 0–8), for 2008 sample A and 2009-2011.
[3] Alternative SDS total score (range 0–4), for 2008 sample B.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2011.

5.6 Evaluation of the Final Weights

5.6.1 Comparisons with the NSDUH Main Adult Sample

To evaluate how well the revised weighting adjustments have addressed bias in the clinical sample, estimates for NSDUH items based on different samples with different weights from 2008 through 2011 are presented in Tables E1 to E3 in Appendix E. The first data column in each of these tables displays estimates from the NSDUH main sample for adults using the NSDUH (main) analysis weights. Tests of statistical significance were carried out between estimates in each of the other data columns and estimates in the first data column. Thus, estimates from the full NSDUH sample are treated as the benchmark for judging the impact of the various weighting adjustments. There is no comparison with estimates based on scaled clinical sample weights because the NSDUH analysis weights are not scaled.

Columns 2, 3, and 4 in Tables E1 to E3 present estimates for adults selected for the clinical interview, nonrespondents (at both phases of nonresponse), and respondents, respectively. All of these estimates use the same weight, MHDSNWT (see Table 5.1 for a description of the weights), which includes the NSDUH adult-level analysis weights, the adjustment for undercoverage of Hispanics[39] who completed the main NSDUH interview in Spanish, and the probability of selection for the clinical interview. Tests of statistical significance were carried out for differences in estimates between nonrespondents (column 3) and respondents (column 4) prior to adjustment for nonresponse.

The final two columns of the tables present estimates for NSDUH items using the revised weights derived from the revised weighting adjustments before scaling (MHNEWWGT) (column 5) and the previous weights derived from the 2008 weighting adjustments (column 6). In addition to comparisons with the "benchmark" estimates from the full NSDUH sample in column 1, the estimates in columns 5 and 6 were also compared with each other to see whether the revised weighting produced any changes in estimates compared with the previous weighting adjustments.

Across all three tables, statistically significant differences between the respondents and nonrespondents in the clinical sample are evident for many items. The descriptive characteristics shown in Table E1 include such demographics as gender, age, race/ethnicity, education, income, region, employment, health insurance coverage, ever having been diagnosed with depression, and country of birth. The mental health items in Table E2 include measures related to K6 scores, WHODAS scores, past year and lifetime measures of MDE, and measures related to suicidal thoughts and mental health treatment. Some of these variables are used in the revised 2012 model for predicting SMI or, at least on the surface, may be correlated with specific disorders assessed on the SCID. The substance use measures shown in Table E3 provide fewer statistically significant differences between respondents and nonrespondents as compared with items shown in Tables E1 and E2.

[39] Tables in Appendix D address the impact of the removal of adults responding to the NSDUH main survey in Spanish from the clinical survey's target population and the effectiveness of weighting to compensate for that.

In addition, the three tables allow comparisons of estimates to be made based on the previous and revised weights. For the largely demographic items in Table E1 and the substance use measures in Table E3, relatively few statistically significant differences can be seen between estimates using the previous and revised weights. Several statistically significant differences are worth noting, however, in the mental health items shown in Table E2. The estimate for lifetime MDE was lower using the revised weights than using the previous weights (12.9 vs. 14.6 percent). Furthermore, the revised weights' MDE estimate was close to the full sample NSDUH estimate (12.8 percent) and was not significantly different. The difference between the estimate using the previous weights (14.6 percent) and the full sample NSDUH estimate, however, was statistically significant. The mean K6 score using the revised weights was 5.1. This was not different (statistically) from the NSDUH full sample estimate of 4.9. However, the mean K6 score of 5.3 under the previous weights was significantly different from the full sample estimate. It should be noted that K6 scores (alternative) were added to the poststratification step for the revised weights.

There appears to be a tendency for the estimates of the respondents prior to adjustment (Table E2, column 4) to more closely resemble the estimates for the final nonrespondents in Table D3b than the initial nonrespondents in Table D2b. For example, the mean K6 score for respondents prior to adjustment was 5.2, while it was 4.9 for final nonrespondents but only 3.2 for initial nonrespondents. Similarly, for lifetime MDE, the estimates were 14.7 percent for respondents prior to adjustment, 12.7 percent for final nonrespondents, and only 5.4 percent for initial nonrespondents.

For a number of measures on mental health treatment and MDE treatment, estimates with the revised weights were to some extent closer to the estimates from the full NSDUH sample than those with the previous weights. These measures were not used in the poststratification adjustment, and most were not used in any of the undercoverage or nonresponse adjustment processes. For example, 6.6 percent of adults aged 18 or older received outpatient mental health treatment in the past 12 months in the full NSDUH sample (Table E2). The estimate for respondents only was 7.7 percent. The revised weights yielded an estimate of 6.9 percent, which was not significantly different from the full NSDUH sample's estimate. In contrast, the previous weights produced an estimate of 7.9 percent, which was significantly different from the estimate for the full NSDUH sample. Other measures in which the estimate produced by the revised weights appeared closer to the full NSDUH sample's estimate than the estimate from the previous weights include receiving any mental health treatment in the past year, using prescription drugs for MDE in the past year, and an unmet need for mental health treatment in the past year.

5.6.2 Unequal Weighting Factors

Another factor to consider in evaluating the revised weighting adjustments is the degree to which the variances of estimates may be affected by the variability of the weights themselves. Of interest here is variation introduced by the implicit use of (sometimes truncated) logistic regression modeling to derive weighting adjustments. When using a modeling approach, adjustment factors may be unstable, which can lead to wide variability in the weights. The unequal weighting effects (UWEs), which are an indicator of the variability of the weights,

are presented in Table 5.8 for each of the weight components examined here. The UWE for year y is as follows:

$$UWE_y = \frac{n_y \sum_{i \in y} w_i^2}{\left(\sum_{i \in y} w_i\right)^2},$$

where w_i represents the weights derived from the revised weighting adjustments before scaling (MHNEWWGT) and n_y is the number of clinical interview respondents in year y.

As a point of comparison, the UWE for the final adult-level weight (ANALWT, or equivalently ANALWT_A) in the 2011 NSDUH for all adults was 2.87. Prior to any adult-level adjustments for the NSDUH main sample (including nonresponse adjustment and poststratification), the UWE was 2.53 (Chen et al., 2014), which implies that the adult-level adjustments for the NSDUH main study inflated the variability of the NSDUH weights. Similarly, the UWE for ANALWT_A increased after the undercoverage adjustment for Hispanics who were interviewed in Spanish (resulted in ANALWT_E) for each year of NSDUH from 2008 to 2012.

Based on the clinical sample, the UWEs for ANALWT_A and ANALWT_E were similar to their counterparts based on the NSDUH main sample, but increased after incorporating the clinical sample selection probability that resulted in the full base weight of the clinical sample (MHDSNWGT). This increase was due to the unequal probability sample design for the clinical sample. Moreover, the UWE for MHDSNWGT generally increased after the first nonresponse adjustment (resulted in MHWTNR1), except for 2012 in which the UWEs for MHDSNWGT and MHWTNR1 were similar (12.04 vs. 11.55). The UWEs for the MHSS clinical sample generally declined from MHWTNR1 to the final analysis weight prior to rescaling (MHNEWWGT). One exception occurred in 2012 when the UWE for the second nonresponse adjustment (MHWTNR2) was a little higher than the UWE for MHWTNR1 (3.74 vs. 3.67). The results suggest that the second nonresponse adjustment and final poststratification were able to smooth the first nonresponse adjustment weights and reduce the weight variability. Also, some of the changes over time in the UWE for each weight component can be accounted for by changes in the sample allocation. For example, the UWE sharply decreased between the 2009 and 2010 surveys, which resulted from a change in the sampling algorithm in which adults in older age groups were oversampled and those in younger age groups were undersampled, which roughly reversed the oversampling in the NSDUH main survey.

5.6.3 Comparisons Based on Standard Error Computations

This section looks at the cumulative impact of the weight adjustments on the SEs of direct estimates computed from the clinical subsample. This analysis incorporates not only the impact on SEs resulting from the UWEs, but also from the poststratification. The former tends to increase SE, while the latter decreases SE. Because the focus is only on the SEs, whatever impact the weight adjustments may have on bias is not considered in the analysis.

Table 5.8 Unequal Weighting Effects for Weights at Stages of Adjustment, by Year of Data Collection: 2008 to 2012 Adult NSDUH Main Study and Clinical Sample

	Main Study		Clinical Sample					
Year/Sample	NSDUH Adult-Level Analysis Weight (ANALWT_A)	Undercoverage Adjustment for Hispanics Interviewed in Spanish (ANALWT_E)	NSDUH Adult-Level Analysis Weight (ANALWT_A)	Undercoverage Adjustment for Hispanics Interviewed in Spanish (ANALWT_E)	Full Base Weight for Clinical Sample (MHDSNWT)	Nonresponse Adjustment for Initial Nonresponse (MHWTNR1)	Nonresponse Adjustment for Final Nonresponse (MHWTNR2)	Analysis Weight for Clinical Sample, Prior to Rescaling (MHNEWWGT)
2008A Count	22,622	21,598	1,194	1,194	1,194	1,024	759	759
2008A UWE	3.15	3.61	3.17	3.28	10.64	13.00	11.11	10.96
2008B Count	23,046	21,991	1,137	1,137	1,137	996	741	741
2008B UWE	3.00	3.92	3.16	3.57	9.94	11.78	9.45	10.46
2009 Count	45,609	43,710	789	789	789	686	520	520
2009 UWE	2.92	3.45	2.73	4.48	12.04	11.55	8.24	7.46
2010 Count	45,844	43,961	768	768	768	644	516	516
2010 UWE	2.98	3.80	2.30	3.39	4.78	5.24	3.94	3.36
2011 Count	46,599	44,734	2,277	2,277	2,277	1,921	1,495	1,495
2011 UWE	2.87	3.52	2.29	3.12	4.06	4.12	3.73	3.07
2012 Count	45,836	44,228	2,464	2,462[1]	2,462[1]	2,062	1,622	1,622
2012 UWE	2.88	3.41	2.11	2.69	3.39	3.67	3.74	3.44

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; UWE = unequal weighting effect.

[1] Two Hispanics selected for the MHSS clinical interview in 2012 had responded to part of the NSDUH main interview in Spanish. As a result, they were treated as if they had been ineligible for the clinical interview when weighting the data (i.e., they were assigned ANALWT_E values of 0).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Using the scaled weights described in Section 5.5.2, Table 5.9 displays direct estimates from the clinical sample for SMI, AMI, and past year MDE using both the final weights (MHFNLWGT) and the "unadjusted" weights (ANALWT_A × MHWT1). The latter treat all sources of nonrandom missingness—whether due to nonresponse or being excluded from the target population—as if they could be ignored.

Table 5.9 Direct Estimates Based on SCID Sample: 2008-2012 MHSS

Variable	Final Weights		Unadjusted Weights[1]		Difference	
	Mean	SE[2]	Mean	SE	Mean	SE
SMI	3.94	0.24	4.13	0.26	-0.19	-0.02
AMI	17.95	0.66	18.36	0.81	-0.40	-0.14
MDE	6.32	0.36	7.12	0.56	-0.79	-0.20

AMI = any mental illness; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP); SE = standard error; SMI = serious mental illness.

NOTE: Estimates in this table were combined across years after scaling the weights as described later in the section.

[1] Estimated means and SEs computed using as the product of the NSDUH main study weight and the inverse of the clinical selection probabilities as the weight. This assumes that missingness whether due to undercoverage or nonresponse was at random.

[2] Computed by applying WTADJX to the combined 2008-2012 data, using scaled versions of w_k and η_k from footnote 38 and separate calibration (CALVARS) and model variables for each year (2008A and 2008B were treated as distinct years).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

One can see that the net impact of the weighting adjustment is that the estimates are lower. The table also computes SEs for both sets of estimates under the assumption that they are not subject to systematic biases (i.e., that the adjustments made to remove potential biases due to the undercoverage of Hispanics and clinical sample nonresponse were unnecessary; this assumption is only required for the unadjusted weights). The SE measures for the estimates computed using the adjusted weights properly account for the impact of the poststratification step,[40] but not fully for the other weighting adjustment steps. That is to say, it accounts for the impact of the steps on the UWE, but not for any further reduction of SE that forcing equations (5.2), (5.4), or (5.6) to hold might cause.

5.7 Summary

Weights are created and used in estimation in an attempt to remove the potential for systematic biases in sample survey estimates. Ideally, a survey respondent's weight is the inverse of his or her probability of selection into the sample, which is known to the statistician drawing the sample. In practice, this weight often needs to be adjusted to account for unit (whole-record) nonresponse to the survey and/or errors in the frame from which the sample was drawn.

[40] WTADJX does this by treating the weighted (by ω_k) mean of yk in the clinical sample as the weighted (by w_k) mean of $\mathbf{z}_k^T\mathbf{b} + (\omega_k/w_k)(y_k - \mathbf{z}_k^T\mathbf{b})$ in the NSDUH main sample, where $\omega_k = \eta_k A_{4bk} = F_k A_{2ak} A_{2bk} A_{3ak} A_{3bk} A_{4ak} A_{4bk}$ is the final clinical weight, $\mathbf{b} = (\sum_S \eta_k A_{4bk}' a_k \mathbf{z}_k \mathbf{z}_k^T)^{-1} \sum_S \eta_k A_{4bk}' a_k \mathbf{z}_k y_k$, and A_{4bk}' is the derivative of the right-hand side of equation (5.7) with respect to $a_k \mathbf{z}_k^T \mathbf{g}$. Kott (2011) pointed out that $\eta_k A_{4bk}'$ is asymptotically identical to 1 and η_k is asymptotically identical to ω_k.

The NSDUH main survey person-level weight incorporated adjustments as well as probabilities of selection. The clinical follow-up survey of adults also required additional adjustments: for the exclusion of adult NSDUH respondents who completed the main interview in Spanish, for those adults who were selected for the clinical follow-up interview but refused to be recontacted, and for those adults who agreed to be recontacted but failed to complete the clinical interview.

The clinical study weighting scheme made separate adjustment for each of these reasons and added a fourth adjustment to reduce the SEs of estimates derived from the clinical sample. A scaling factor was applied to the revised weights in each year before combining the data across years because the clinical interview sample sizes, as well as the sample allocations and weighting adjustments, varied from year to year. Applying this scaling reduced the variance of estimates for mental health characteristics estimated directly from the clinical sample and was intended to increase the efficiency of the modeling that is described in Chapter 7.

6. MHSS Clinical Sample Characteristics

6.1 Background

Both the annual weighted and unweighted Mental Health Surveillance Study (MHSS) clinical interview samples were evaluated each year from 2008 to 2012. The unweighted samples were evaluated to ensure that the sample composition reflected the sampling algorithm (e.g., the oversampling of higher Kessler-6 [K6] scores in 2008). The weighted samples were evaluated to ensure that the weighted distribution of the demographic and other characteristics of the clinical sample was similar to the National Survey on Drug Use and Health (NSDUH) main interview data.

This evaluation was done by comparing demographic and key mental health and substance use characteristics in the clinical sample across years. This chapter provides a summary of the single-year evaluations that were conducted and a description of the combined 2008 to 2012 clinical data.[41]

6.2 Summary of Prior Sample Evaluations

As described in Chapter 3, a split-sample design was implemented in 2008 for NSDUH in which a random half of the sample received an abbreviated version of the World Health Organization Disability Assessment Schedule (WHODAS) and the other half received the Sheehan Disability Scale (SDS). The randomization of the impairment scales was maintained within this clinical interview subsample, so that about half of the clinical interview participants were administered the WHODAS and half were administered the SDS. Therefore, the 2008 MHSS clinical sample evaluation consisted of descriptive analyses and statistical tests that were conducted to examine the distribution of respondent characteristics in the full clinical sample and to compare the two half samples, The purpose of these analyses was to determine whether estimates from the two half samples were comparable without accounting for differences caused by random sampling between the two samples. Key demographic characteristics examined included gender, age, race/ethnicity, and education. Substance use characteristics included past month tobacco and marijuana use and past year alcohol abuse or dependence. The mental health characteristics included Structured Clinical Interview for DSM-IV (SCID) diagnoses, serious mental illness (SMI) status, and K6 scores. Further details about these variables can be obtained from the 2008 NSDUH national findings report (Office of Applied Studies [OAS], 2009).

Analyzing both the unweighted and weighted data from the 2008 MHSS clinical study indicated that none of the demographic or mental health measures were significantly different across the two samples. There were, however, some large differences in the substance use measures, including past month tobacco and marijuana use. For details, see Aldworth et al. (2009).

[41] The weights used in the prior descriptive analyses were the old MHSS analysis weights (e.g., Liao et al., 2012), whereas the 2012 descriptive analyses used the new MHSS analysis weights developed in 2013. Chapter 5 provides more details about the components of these new weights.

For the 2009 MHSS clinical sample evaluation, initial descriptive analyses and statistical tests compared key demographic and psychosocial characteristics between the 2009 MHSS clinical sample and the 2008 MHSS WHODAS clinical sample. The purpose of these analyses was to determine whether the 2009 sample characteristics (see Section 3.5 for details) were consistent with the data collection from the WHODAS half sample of 2008. Key demographic characteristics included gender, age, race/ethnicity, and education. Mental health characteristics included lifetime and past year depression, depression treatment, mental health treatment, and suicidality measures. Substance use characteristics included past month tobacco and marijuana use and past year alcohol abuse or dependence. Additional mental health characteristics from the SCID (i.e., the MHSS clinical interview) included SMI status and diagnoses of mental health and substance use. Further details about these variables can be obtained from the 2009 NSDUH mental health findings report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2010) and the 2009 NSDUH national findings report (OAS, 2010a, 2010b).

None of the unweighted demographic characteristics were significantly different between the two samples (i.e., the 2009 sample and the WHODAS half sample of 2008), but some of the prevalence estimates of mental health and substance use measures differed significantly. After applying the weights, most of these differences were no longer statistically significant. There were, however statistically significant (i.e., $p < 0.05$) and marginally significant (i.e., $p < 0.10$) differences, respectively, for past month cigarette use and past year alcohol abuse and dependence between the two samples. Differences in the past year K6 total scores were statistically significant before the weights were applied, but they were no longer significantly different once the data were weighted. For details, see Aldworth et al. (2010).

For the 2010 MHSS clinical sample evaluation, initial descriptive analyses and statistical tests compared key demographic and psychosocial characteristics in the clinical sample that were collected in 2008 and the first two quarters of 2009 with the clinical sample collected in the final two quarters of 2009 and the full clinical sample from 2010. The purpose of this particular comparison was to determine whether the expected changes based on the revisions in sampling allocation design after the second quarter of 2009 were being realized (see Sections 3.5 and 3.6 for details) and whether the weights were appropriately accounting for these changes. Further changes to the sampling design were made in 2010. To determine whether changes in the sampling design affected the demographic, mental health, and substance use characteristics between the samples, comparisons of these measures were made between the final two quarters of 2009 and 2010. Key demographic characteristics included gender, age, race/ethnicity, and education. Mental health characteristics included lifetime and past year depression, depression treatment, mental health treatment, and suicidality measures. Substance use characteristics included past month tobacco and marijuana use and past year alcohol abuse or dependence. Additional mental health characteristics from the SCID included SMI status and diagnoses of mental health and substance use. Details about these variables can be obtained from the 2010 NSDUH national findings report (CBHSQ, 2011) and the 2010 NSDUH mental health findings report (CBHSQ, 2012a).

Between the first two time periods and the final two time periods, unweighted estimates of age and education were significantly different. Some of the prevalence estimates of the mental health and substance use measures also appeared to be significantly different between the first two time periods and the final two time periods. Once the data were weighted, however, there

were fewer significant differences. Specifically, only the difference of past year alcohol and illicit drug abuse or dependence was still statistically significant ($p < 0.05$). The K6 scores within the SCID cases (i.e., the MHSS clinical interview cases) were compared between the two samples by Cochran-Mantel-Haenszel (CMH) tests of general association (24 *df* [degrees of freedom]) and means (1 *df*). For the unweighted scores, the two tests were statistically significant, but for the weighted scores only the test of general association was statistically significant. For details, see Aldworth et al. (2012).

For the 2011 MHSS clinical sample evaluation, initial descriptive analyses and statistical tests compared key demographic and psychosocial characteristics between the WHODAS sample of 2008-2010 with the WHODAS sample of 2011. The purpose of this particular comparison was to assess the consistency of the sample collected in 2011 with the samples collected from prior years. Two further comparisons were also made. Estimates for 2008-2009 were compared with estimates for 2010-2011 to determine whether changes in the sampling design implemented in 2010 and 2011 (see Section 3.6 for details) affected the demographic, mental health, and substance use characteristics between the samples. In addition, estimates for 2010 were compared with those for 2011 to examine the impact of random sampling variation occurring under the same sampling design with an increased sample size. Key demographic characteristics included gender, age, race/ethnicity, education, poverty threshold, core-based statistical area (CBSA), and employment status. Mental health characteristics included lifetime and past year depression, depression treatment, mental health treatment, and suicidality measures. Substance use characteristics included past month tobacco and marijuana use and past year alcohol abuse or dependence. Mental health characteristics from the SCID included SMI status and diagnoses of mental health and substance use. Details about these variables can be obtained from the 2011 NSDUH national findings report (CBHSQ, 2012b) and the 2011 NSDUH mental health findings report (CBHSQ, 2012c).

Overall, there appears to be evidence that modifications to the sampling design from 2008 to 2011 had differential effects on the unweighted distributions of several key characteristics in each year's MHSS clinical data. The modifications to the sampling design reduced the oversampling of respondents with high K6 scores and increased the undersampling of respondents with low K6 scores. This had the effect that the sampled distributions were more similar to the population distributions, and the variation in the size of the weights was also reduced. As proven through the weighted descriptive analysis, taking weights into account removed most of these differential effects. However, there were still some significant effects for the characteristics in the weighted data, including the following:

- education when comparing 2008-2010 with 2011;
- past year receiving outpatient mental health treatment when comparing 2008-2010 with 2011, when comparing 2008-2009 with 2010-2011, and when comparing 2010 with 2011;
- past year receiving any mental health treatment when comparing 2008-2010 with 2011;
- Global Assessment of Functioning (GAF) score less than or equal to 50 when comparing 2008-2010 with 2011; and

- any mental illness (AMI) when comparing 2008-2010 with 2011.

When making comparisons of the estimates across different time periods, attention should be paid to whether the compared estimates are correlated with one or more characteristics that have imbalanced distributions across the compared time periods. Differences in estimates across time periods may not reflect a true difference in populations across the compared time periods, but result instead from the fact that the distributions of related characteristics are different (i.e., are imbalanced) in the samples collected from the different time periods. For details, see Liao et al. (2012).

6.3 Evaluating the Overall 2008 to 2012 Sample

Initial descriptive analyses and statistical tests compared key demographic, mental health, and substance use characteristics between the MHSS clinical sample from each individual year and the combined MHSS clinical samples from 2008 through 2012. The purpose of this particular comparison was to assess the consistency of samples collected from different sampling periods.

Key demographic characteristics included age, gender, race/ethnicity, education, poverty threshold, CBSA, and employment status. Mental health characteristics included lifetime and past year depression, depression treatment, mental health service use, suicidality measures, and some mental health characteristics from the SCID, including mental illness status and substance use disorder. Substance use characteristics included past month substance use and past year substance abuse or dependence. Details about these variables can be obtained from the 2011 NSDUH national findings report (CBHSQ, 2012b) and the 2011 NSDUH mental health findings report (CBHSQ, 2012c).

Unweighted descriptive statistics of the demographic, mental health, and substance use measures based on the MHSS clinical sample are shown in Tables 6.1, 6.2, and 6.3, and weighted versions of those descriptive statistics based on both the MHSS clinical sample and the all-adult NSDUH main interview sample are shown in Tables 6.4, 6.5, and 6.6.[42] Included in the descriptive statistics are percentages across the following six sampling periods: (1) 2008, (2) 2009, (3) 2010, (4) 2011, (5) 2012, and (6) the combined 2008-2012. The weighted percentages for the combined 2008-2012 clinical sample used the scaled analysis weights.[43] For comparison purposes, the combined 2008-2012 adult NSDUH main study samples used both the unscaled NSDUH analysis weights and the scaled NSDUH analysis weights with the same scaling factors as were applied to the clinical sample. Statistical tests compared each sampling period from the clinical sample and the adult NSDUH main study sample, as well as the combined 2008-2012 sampling period from the adult NSDUH main study sample with the combined 2008-2012 sampling period from the clinical sample.

[42] To facilitate the data presentation and discussion, all of the tables in this chapter have been grouped at the end of the chapter's text.

[43] Scaling was applied to the sample weights to reduce the variance of estimates and increase statistical efficiency. See Section 5.5 for more details.

6.3.1 Demographic Characteristics

Table 6.1 shows that with unweighted data there was a statistically significant difference between some individual year clinical sample and the combined 2008-2012 clinical sample for some demographic characteristics, including age, education, poverty, population density, and employment status. These differences appeared to be driven by both the initial change in the sampling design after the first two quarters of 2009 and the change in 2010, which allowed a greater proportion of older respondents and persons with lower K6 and WHODAS scores to be sampled for the MHSS clinical study. No significant effect was evident in the tests for gender and Hispanic origin and race.

Table 6.4 indicates that applying weights mitigated most of the differences among the demographic characteristics that appeared in the unweighted data of the clinical sample. That is, corresponding *p* values increased and often became insignificant. The only significant difference remaining was for adults with an education level of less than high school in the 2008 clinical sample. Note that although the 2008-2012 unweighted proportion of this subpopulation was lower than its corresponding 2008 counterpart, the 2008-2012 weighted proportion was higher than its corresponding 2008 counterpart. However, three insignificant differences among the unweighted statistics turned into significant differences after the weights were applied: (1) adults living in a CBSA with 1 million or more people in 2008, (2) adults who were unemployed in the past year in 2008, and (3) adults who were in the "other" category of employment status in 2009. Among these three differences, the weighted proportions of the combined 2008-2012 clinical sample were all relatively higher than their corresponding counterparts in 2008 or 2009.

Table 6.4 also displays weighted percentages derived from the main computer-assisted interviewing (CAI) sample (i.e., the all-adult NSDUH main interview sample). As shown in the last two columns, the 2008-2012 CAI sample distributions based on both the unscaled and scaled weights were generally consistent with their corresponding SCID counterparts without any significant differences, which implies that the revised weighting procedures for the clinical sample (as described in Chapter 5) effectively adjusted the clinical sample distributions of these key demographic characteristics to be consistent with the all-adult NSDUH main study sample distributions. There was only one exception for adults with a poverty level larger than or equal to the 200 percent threshold. The weighted percentage for this group in the CAI sample with scaled weights was slightly smaller than its counterpart in the SCID sample (65.9 vs. 68.6 percent, respectively). A few differences were also found when comparing the 2008-2012 SCID weighted percentages with the individual year CAI weighted percentages. These differences had less of an impact in these analyses because scaled weights were used for analyzing the clinical sample, while unscaled weights were used for analyzing the all-adult NSDUH main study sample.

6.3.2 Mental Health Characteristics

Table 6.2 presents the unweighted percentages of some mental health characteristics in the SCID clinical sample. As shown in the table, significant differences were found in all four mental health categories when comparing an individual year sample with the combined 2008-2012 sample, including major depressive episode (MDE), past year treatment for depression, past year mental health service use, suicidal experiences, and some key variables from the SCID.

After the weights were applied, as shown in Table 6.5, all but one of the significant differences in the unweighted data became insignificant. The only significant difference remaining was for the adults who received inpatient mental health service in the past year in 2011. Also, three insignificant differences in the unweighted data became significant when the weights were applied: (1) adults who received inpatient mental health service in the past year in 2009, (2) adults who received outpatient mental health service in the past year in 2011, and (3) adults who had attempted suicide in the past year in 2010. Although these results may be due to real underlying differences in the prevalence rates being estimated, they also can be due to a combination of the small yearly clinical sample sizes and the small prevalence rates. Tests performed for differences assume the asymptotic normality of the test statistic. This assumption may not be reasonable when there are small sample sizes and prevalence rates. As in Table 6.4, Table 6.5 also displays weighted percentages derived from the CAI sample. No significant difference was found between the 2008-2012 SCID sample and the 2008-2012 CAI sample, which implies that the sample distributions of these two samples were consistent for the compared key characteristics in this table.

6.3.3 Substance Use Characteristics

Table 6.3 indicates that for the unweighted clinical sample, there was a significant difference between some of the individual samples (including clinical samples from 2008, 2009, 2011, and 2012 and the combined 2008-2012 clinical sample) for some substance use characteristics, including past month substance use and past year substance abuse or dependence. However, there were no significant differences between the 2010 clinical estimates and combined 2008-2012 clinical sample. Table 6.6 shows that after applying weights to the data there were no longer any significant differences between any individual year clinical sample estimates and the combined 2008-2012 clinical sample estimates. The weighted distributions of all of these substance use characteristics were consistent among the SCID and CAI samples across different sampling periods.

6.3.4 K6 and WHODAS Total Scores

Unweighted descriptive statistics of the past year K6 total score (i.e., the maximum of the past 30-day K6 total score and the worst month K6 total score) of the SCID samples are shown in Table 6.7, and similar weighted descriptive statistics of the SCID samples and the CAI samples are shown in Table 6.8. The K6 scores were compared between the individual year SCID samples and the CAI samples with the combined 2008-2012 SCID sample by CMH tests of general association (24 *df*) and means (1 *df*). For the unweighted scores, results show that each individual year SCID sample was significantly different from the combined 2008-2012 SCID sample, except for the 2010 SCID sample with respect to the test of means. Weights removed all of the other significant differences for the tests of means. However, all of the general association tests still remained significant for the weighted scores in the SCID sample, which might be caused by sampling design changes on the distribution of past year K6 total scores across years. Both individual year CAI samples and the combined 2008-2012 CAI sample with unscaled or scaled weights showed no differences when compared with the combined 2008-2012 SCID sample with respect to the tests of means. Similar to the individual year SCID samples, all of the general association tests were statistically significant for the weighted scores in the CAI samples when compared with the combined 2008-2012 SCID sample. Weighted descriptive statistics of

the past 30-day K6 total score and the past year K6 total score for all of the sampling periods are given in Table 6.9. Their means were compared between the individual year SCID samples and the combined 2008-2012 SCID sample by the *t* test. All of the means from each individual year SCID sample were found insignificantly different from the means of the combined 2008-2012 SCID sample. In addition, no difference was found for the 2008-2012 CAI sample with unscaled or scaled weights when compared with the 2008-2012 SCID sample.

The weighted frequency distributions of the past year WHODAS total scores for the SCID and CAI samples are given in Table 6.10. The WHODAS scores were also compared between the individual year SCID samples, the individual year CAI samples, and the combined 2008-2012 CAI sample with the combined 2008-2012 SCID sample, respectively, by CMH tests of general association (24 *df*) and means (1 *df*). Similar to the results for the K6 total score, no significant difference was evident among the tests of means with the weighted data, while significant differences were detected among all of the tests of general association except for the 2009 CAI sample.

6.3.5 Summary

Overall, some significant differences were found for some key demographic, mental health, and substance use characteristics when comparing each individual year clinical sample with the combined 2008-2012 clinical sample, which indicates some inconsistencies across unweighted clinical samples from different sampling periods. These differences could be driven by four factors: (1) the initial change in the sampling design after the first two quarters of 2009 and the change in 2010, (2) the small yearly clinical sample, (3) the small yearly prevalence rates of the characteristics, and (4) the real underlying changes of the yearly prevalence rates across different sampling periods in the clinical study.

Applying weights to the data removed most of these differential effects between estimates. However, there were still some significant differences for the characteristics in the weighted data, as follows:

- education level less than high school (in 2008),
- living in a CBSA with 1 million or more people (in 2008),
- unemployed in the past year (in 2008),
- having "other" employment status (in 2009),
- received inpatient mental health service in the past year (in 2011),
- received inpatient mental health service in the past year (in 2009),
- received outpatient mental health service in the past year (in 2011), and
- attempted suicide in the past year (in 2010).

When making comparisons of the estimates across different sampling periods, attention should be paid to whether the compared estimates are correlated with one or more characteristics that have imbalanced distributions across the compared time periods. Differences in estimates across time periods may result from different sample distributions of these characteristics and

thus cannot reflect a true difference in populations across the compared time periods. In addition, no significant test was found between the weighted 2008-2012 clinical sample (with scaled weights) and the weighted 2008-2012 all-adult NSDUH main interview sample (with unscaled weights), which implies that the revised MHSS analysis weights have adjusted the clinical sample well enough to be consistent with the all-adult NSDUH interview study sample.

Table 6.1 Sample Description of SCID Respondents, by Demographic Characteristics: 2008-2012

Characteristic	2008 Unweighted Pct.	2009 Unweighted Pct.	2010 Unweighted Pct.	2011 Unweighted Pct.	2012 Unweighted Pct.	2008-2012 Unweighted Pct.
Total (Sample Size)	1,500	520	516	1,495	1,622	5,653
Age						
18-25	58.7[a]	56.3[a]	24.8[a]	20.9[a]	21.3[a]	34.7
26-34	14.9[a]	14.2[a]	25.8[a]	24.7[a]	22.3	20.6
35-49	17.9[a]	18.8[a]	33.7[a]	31.0[a]	30.1[a]	26.4
50+	8.5[a]	10.6[a]	15.7	23.4[a]	26.3[a]	18.4
Age						
18-30	67.2[a]	63.8[a]	40.3[a]	34.0[a]	33.9[a]	46.1
31+	32.8[a]	36.2[a]	59.7[a]	66.0[a]	66.1[a]	53.9
Gender						
Male	36.5[a]	42.1	40.3	40.5	38.9	39.1
Female	63.5[a]	57.9	59.7	59.5	61.1	60.9
Hispanic Origin and Race						
Not Hispanic or Latino						
White	71.6	72.3	72.1	73.2	72.7	72.5
Black or African American	10.8	10.0	10.3	9.6	10.2	10.2
Other or Multiple Races	7.5	7.1	8.1	7.4	7.6	7.5
Hispanic or Latino	10.1	10.6	9.5	9.8	9.4	9.8
Education						
< High School	14.3[a]	11.7	10.5	9.6[a]	10.3	11.3
High School Graduate	29.3	31.3	26.9	27.3	27.9	28.3
Some College	33.4	31.5	32.0	32.6	30.1[a]	31.9
College Graduate	23.1[a]	25.4	30.6	30.5[a]	31.7[a]	28.4
Poverty[1]						
< 100% Threshold	18.1	17.7	11.9[a]	15.2	16.8	16.3
100-199% Threshold	23.4	21.7	21.6	22.1	20.6	21.9
≥ 200% Threshold	58.4[a]	60.6	66.5[a]	62.8	62.6	61.7
Population Density						
CBSA ≥ 1M	39.1	40.2	43.4	38.3	40.4	39.7
250K ≤ CBSA < 1M	25.5	26.0	23.3	26.0	22.7[a]	24.7
CBSA < 250K	26.3	27.3	24.0	26.7	28.7	27.0
Non-CBSA, Urban	3.6	2.1	2.7	2.9	1.7[a]	2.7
Non-CBSA, Rural	5.4	4.4	6.6	6.2	6.5	5.9
Employment Status						
Full Time	49.5	49.2	55.0	49.9	51.9	50.8
Part Time	23.4[a]	21.2	15.1[a]	16.7[a]	16.5[a]	18.7
Unemployed	6.5	8.7	9.3	6.7	6.6	7.0
Other[2]	20.7[a]	21.0	20.5	26.8[a]	25.0	23.5

CAI = computer-assisted interviewing; CBSA = core-based statistical area; K = thousand; K6 = Kessler-6, a 6-item psychological distress scale; M = million; NSDUH = National Survey on Drug Use and Health; pct. = percent; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. The respondent characteristics in this table are based on data collected from the NSDUH CAI interview.

[a] Difference between this estimate and the corresponding estimate from 2008-2012 is statistically significant at the .05 level.

[1] U.S. census poverty threshold. Persons aged 18 to 22 in a college dormitory were excluded from the analysis.

[2] The other employment category includes students, persons keeping house or caring for children full time, retired or disabled persons, or other persons not in the labor force.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 6.2 Sample Description of SCID Respondents, by Mental Health Characteristics: 2008-2012

Characteristic	2008 Unweighted Pct.	2009 Unweighted Pct.	2010 Unweighted Pct.	2011 Unweighted Pct.	2012 Unweighted Pct.	2008-2012 Unweighted Pct.
Total (Sample Size)	1,500	520	516	1,495	1,622	5,653
MDE[1]						
Lifetime/Not Past Year MDE	11.5	11.1	12.7	10.3	10.5	11.0
Past Year MDE						
Without Impairment	7.3[a]	6.6	6.3	5.3	4.8[a]	5.9
With Impairment	14.2	12.0	15.6	12.1[a]	15.2	13.9
No Lifetime/Past Year MDE	66.9[a]	70.3	65.4	72.3[a]	69.5	69.3
Past Year Treatment for Depression[2]						
Any Treatment	63.6[a]	56.3[a]	74.1	73.0	72.1	68.8
Saw or Talked to Health Professional	60.0	51.0[a]	67.9	69.1	67.2	64.3
Prescription Medication	41.3[a]	44.8	64.3[a]	56.8	56.3	52.1
Past Year Mental Health Service Use[3]						
Any Treatment	25.6	25.3	32.2[a]	24.6	27.2	26.4
Outpatient	15.3	16.4	18.1	13.4	15.1	15.1
Inpatient	1.0	0.6	1.7	0.7[a]	1.7[a]	1.2
Prescription Medication	20.7	20.6	27.1[a]	20.5	23.9	22.1
Suicidal Experiences[4]						
Had Thoughts of Suicide	13.7[a]	10.0	10.5	9.7	9.7[a]	10.9
Made Plans for Suicide	4.5[a]	3.1	2.9	2.9	2.9	3.3
Attempted Suicide	2.1[a]	0.8	1.0	1.0	1.3	1.4
SCID Variables[5]						
Any Mental Illness[6]	43.8[a]	44.4[a]	37.4	33.9[a]	33.9[a]	37.8
Serious Mental Illness	11.5	10.2	12.4	11.0	11.4	11.3
Moderate Mental Illness	11.3	10.2	10.9	10.8	11.0	10.9
Mild Mental Illness	21.1[a]	24.0[a]	14.1	12.1[a]	11.5[a]	15.6
No Mental Illness	56.2[a]	55.6[a]	62.6	66.1[a]	66.1[a]	62.2
Substance Use Disorder	16.6[a]	16.2	13.2	10.9[a]	10.1[a]	12.9

CAI = computer-assisted interviewing; GAF = Global Assessment of Functioning; K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; NSDUH = National Survey on Drug Use and Health; pct. = percent; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. Most of the respondent characteristics in this table are based on data collected from the NSDUH CAI interview, except for the mental illness and substance use disorder status. These characteristics were determined based on the SCID data. The SCID was the diagnostic tool used for the clinical interview follow-up study to the main NSDUH study.

[a] Difference between this estimate and the corresponding estimate from 2008-2012 is statistically significant at the .05 level.

[1] Respondents with unknown past year MDE data were excluded.
[2] Among those with MDE. "Any Treatment" includes categories either "Saw or Talked to Health Professional" or "Prescription Medication."
[3] Mental health treatment/counseling is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.
[4] Respondents with unknown suicide information were excluded.
[5] Variables from this section were based on the SCID clinical interview, while other characteristic variables in Tables 6.1 to 6.10 were based on the NSDUH main interview.
[6] Three categories of the level of mental illness severity based on functional impairment are defined based on the SCID disorder diagnosis and GAF scales: mild mental illness, moderate mental illness, and serious mental illness. Any mental illness includes persons in any of the three categories. For more details, see Table 1.1 in Chapter 1.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 6.3 Sample Description of SCID Respondents, by Substance Use Measures: 2008-2012

Characteristic	2008 Unweighted Pct.	2009 Unweighted Pct.	2010 Unweighted Pct.	2011 Unweighted Pct.	2012 Unweighted Pct.	2008-2012 Unweighted Pct.
Total (Sample Size)	1,500	520	516	1,495	1,622	5,653
Past Month Substance Use						
Illicit Drugs[1]	15.9[a]	18.8[a]	15.1	13.0	10.8[a]	13.9
Marijuana	13.3[a]	15.0[a]	11.4	10.5	8.7[a]	11.2
Illicit Drugs Other Than Marijuana[1]	6.7[a]	7.9[a]	6.6	4.2[a]	3.7[a]	5.3
Tobacco Products[2]	40.3[a]	41.0	38.6	35.5	34.3[a]	37.2
Alcohol	61.9	65.0	62.8	60.7	61.7	61.9
Past Year Substance Abuse or Dependence						
Substance Use Disorder[3]	22.3[a]	20.8[a]	13.6	13.4[a]	13.8[a]	16.6
Alcohol Use Disorder	18.5[a]	17.7[a]	12.0	10.5[a]	11.2[a]	13.6
Illicit Drug Use Disorder	7.9[a]	7.7	4.5	4.7[a]	4.6[a]	5.8
No Substance Use Disorder	77.7[a]	79.2[a]	86.4	86.6[a]	86.2[a]	83.4

CAI = computer-assisted interviewing; K6 = Kessler-6, a 6-item psychological distress scale; NSDUH = National Survey on Drug Use and Health; pct. = percent; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. The respondent characteristics in this table are based on data collected from the NSDUH CAI interview.

[a] Difference between this estimate and the corresponding estimate from 2008-2012 is statistically significant at the .05 level.

[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other Than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically, including data from original methamphetamine questions but not including new methamphetamine items added in 2005 and 2006.

[2] Tobacco Products include cigarettes, smokeless tobacco (i.e., chewing tobacco or snuff), cigars, or pipe tobacco.

[3] Substance Use Disorder is defined as meeting criteria for illicit drug or alcohol dependence or abuse. Dependence or abuse is based on definitions found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 6.4 Weighted Sample Description of SCID and CAI Respondents, by Demographic Characteristics: 2008-2012

Characteristic	2008		2009		2010		2011		2012		2008-2012		
	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	
	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct. (unscaled)	Wt. Pct. (scaled)
Total (Numbers in Thousands)	224,923	224,923	227,207	227,207	229,273	229,273	232,625	232,625	235,124	235,124	231,890	229,830	231,890
Age													
18-25	14.6	14.6	14.8	14.8	14.9	14.9	14.7	14.7	14.7	14.7	14.7	14.7	14.7
26-34	18.6	15.8	19.1	15.9	14.5	15.9	16.6	15.7	16.6	15.7	16.6	15.8	15.7
35-49	25.8	28.5[a]	24.7	27.8	28.5	27.1	25.6	26.5	25.1	26.0	25.8	27.2	26.7
50+	41.0	41.0	41.5	41.5	42.1	42.1	43.1	43.1	43.6	43.6	42.8	42.3	42.8
Age													
18-30	25.6	23.9	26.3	23.8	24.3	24.0	23.6	23.7	24.0	23.7	24.2	23.8	23.8
31+	74.4	76.1	73.7	76.2	75.7	76.0	76.4	76.3	76.0	76.3	75.8	76.2	76.2
Gender													
Male	48.3	48.3	48.3	48.3	48.4	48.4	48.1	48.1	48.1	48.1	48.2	48.2	48.2
Female	51.7	51.7	51.7	51.7	51.6	51.6	51.9	51.9	51.9	51.9	51.8	51.8	51.8
Hispanic Origin and Race													
Not Hispanic or Latino													
White	68.8	68.8	68.4	68.4	68.0	68.0	66.7	66.7	66.3	66.3	67.1	67.6	67.1
Black or African American	11.3	11.3	11.6	11.6	11.6	11.6	11.5	11.5	11.6	11.6	11.5	11.5	11.5
Other or Multiple Races	6.4	6.4	6.3	6.3	6.4	6.4	7.2	7.2	7.4	7.4	7.0	6.7	7.0
Hispanic or Latino	13.5	13.5	13.7	13.7	13.9	13.9	14.6	14.6	14.8	14.8	14.4	14.1	14.4
Education													
< High School	7.6[a]	15.6[a]	14.2	15.4	13.8	15.3[a]	12.8	14.2	14.8	14.6	13.1	15.0	14.7
High School Graduate	32.3	31.0	28.1	30.7	31.3	30.2	28.3	30.0	28.8	29.7	29.4	30.3	30.0
Some College	30.3	25.5	23.4	25.4	25.9	25.8	29.7	26.5	25.2	26.6	27.4	26.0	26.3
College Graduate	29.8	27.9	34.3	28.6	29.1	28.7	29.3	29.3	31.2	29.2	30.2	28.7	29.0
Poverty[1]													
< 100% Threshold	11.9	11.3	9.3	12.2	9.5	13.0	13.2	14.1	13.3	15.9[a]	12.4	13.3	14.2
100-199% Threshold	14.3	18.6	19.6	19.4	16.9	20.8	19.4	20.6	20.7	19.4	19.0	19.8	19.9
≥ 200% Threshold	73.7	70.1	71.1	68.3	73.6	66.2	67.3	65.3[a]	66.0	64.8[a]	68.6	66.9	65.9[a]
Population Density													
CBSA ≥ 1M	42.7[a]	51.9	50.1	51.7	56.0	52.1	50.0	51.7	52.1	51.9	50.7	51.8	51.8
250K ≤ CBSA < 1M	24.7	21.4	23.7	21.5	17.9	21.7	24.7	22.6	20.1	21.9	22.1	21.8	22.0
CBSA < 250K	23.1	20.4	17.5	20.8	19.5	20.0	19.7	19.5	21.8	20.0	20.7	20.1	19.9
Non-CBSA, Urban	2.8	2.0	0.9	1.8	1.3	2.1[a]	1.6	1.8	1.2	1.6	1.5	1.8	1.8
Non-CBSA, Rural	6.8	4.3	7.8	4.2	5.2	4.2	4.0	4.6	4.8	4.7	4.9	4.4	4.5
Employment Status													
Full Time	57.8	54.4[a]	60.9	50.5	53.0	49.8	47.3	49.7	50.8	50.0	51.1	50.9	50.4
Part Time	12.6	13.5	14.6	14.0	11.9	14.5	13.7	13.9	13.2	13.9	13.2	14.0	14.0
Unemployed	2.6[a]	4.0[a]	5.2	6.6	8.7	6.5	7.7	5.8	6.0	5.8	6.5	5.7	5.7
Other[2]	27.0	28.2	19.3[a]	29.0	26.4	29.3	31.4	30.5	30.0	30.3	29.2	29.4	29.9

CAI = computer-assisted interviewing; CBSA = core-based statistical area; K = thousand; K6 = Kessler-6, a 6-item psychological distress scale; M = million; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule; wt. pct. = weighted percent.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. The respondent characteristics in this table are based on data collected from the NSDUH CAI interview.

NOTE: For each individual year, the MHSS or the NSDUH annual analysis weight was used to obtain the weighted estimates. For 2008-2012, the rescaled MHSS, the scaled NSDUH, or the unscaled NSDUH analysis weight for the combined 5-year data was used to obtain the weighted estimates. For 2008, MHSAMPWT was used to obtain the NSDUH estimates for the single year data as well as for the combined 5-year data.

[a] Difference between this estimate and the corresponding estimate from the 2008-2012 SCID data is statistically significant at the .05 level.

[1] U.S. census poverty threshold. Persons aged 18 to 22 in a college dormitory were excluded from the analysis.

[2] The other employment category includes students, persons keeping house or caring for children full time, retired or disabled persons, or other persons not in the labor force.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table 6.5 Weighted Sample Description of SCID and CAI Respondents, by Mental Health Characteristics: 2008-2012

Characteristic	2008		2009		2010		2011		2012		2008-2012		
	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	
	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct. (unscaled)	Wt. Pct. (scaled)
Total (Numbers in Thousands)	224,923	224,923	227,207	227,207	229,273	229,273	232,625	2 32,625	235,124	235,124	231,890	229,830	231,890
MDE[1]													
Lifetime/Not Past Year MDE	6.4	6.4	6.5	6.4	5.8	5.9	6.1	6.1	6.3	6.3	6.2	6.2	6.2
Past Year MDE													
Without Impairment	2.4	2.4	3.0	2.6	2.2	2.6	2.8	2.4	1.8	2.4	2.3	2.5	2.4
With Impairment	4.4	4.4	3.7	4.0	4.7	4.2	3.8	4.2	5.0	4.5	4.4	4.3	4.3
No Lifetime/Past Year MDE	86.8	86.8	86.8	87.0	87.4	87.3	87.3	87.3	86.8	86.8	87.1	87.0	87.1
Past Year Treatment for Depression[2]													
Any Treatment	70.5	69.0	76.4	64.3	67.6	68.2	66.8	68.1	70.4	68.0	69.0	67.5	68.0
Saw or Talked to Health Professional	61.4	63.5	75.0	58.9	65.0	62.7	62.2	62.2	66.0	62.7	64.4	62.0	62.5
Prescription Medication	50.5	52.9	65.4	50.0	60.9	52.1	48.3	51.7	53.7	51.6	52.9	51.7	51.8
Past Year Mental Health Service Use[3]													
Any Treatment	15.9	13.4	12.3	13.4	14.7	13.8	12.0	13.6	13.9	14.5	13.5	13.8	13.9
Outpatient	8.1	6.8	7.8	6.4	6.7	6.6	4.9[a]	6.7	6.7	6.6	6.3	6.6	6.6
Inpatient	0.4	0.9	0.2[a]	0.8	1.4	0.8	0.2[a]	0.8	1.0	0.8	0.7	0.8	0.8
Prescription Medication	13.3	11.4	10.9	11.3	11.3	11.7	10.3	11.5	11.6	12.4	11.3	11.7	11.8
Suicidal Experiences[4]													
Had Thoughts of Suicide	3.7	3.7	3.7	3.7	3.8	3.8	3.7	3.7	3.8	3.9	3.8	3.8	3.8
Made Plans for Suicide	1.0	1.0	1.3	1.0	1.3	1.1	0.8	1.0	1.2	1.1	1.0	1.1	1.1
Attempted Suicide	0.2	0.5	0.2	0.5	0.1[a]	0.5	0.3	0.5	0.6	0.6[a]	0.4	0.5	0.5
SCID Variables[5]													
Any Mental Illness[6]	19.8	--	20.3	--	17.7	--	17.1	--	18.1	--	18.0	--	--
Serious Mental Illness	4.7	--	5.6	--	3.2	--	3.7	--	4.1	--	3.9	--	--
Moderate Mental Illness	4.2	--	5.0	--	4.0	--	5.9	--	4.9	--	5.1	--	--
Mild Mental Illness	10.9	--	9.6	--	10.5	--	7.4	--	9.2	--	9.0	--	--
No Mental Illness	80.2	--	79.7	--	82.3	--	82.9	--	81.9	--	82.0	--	--
Substance Use Disorder	7.3	--	6.3	--	9.2	--	7.0	--	8.1	--	7.7	--	--

CAI = computer-assisted interviewing; GAF = Global Assessment of Functioning; K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule; wt. pct. = weighted percent.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. Most of the respondent characteristics in this table are based on data collected from the NSDUH CAI interview, except for the mental illness and substance use disorder status. These characteristics were determined based on the SCID data. The SCID was the diagnostic tool used for the clinical interview follow-up study to the main NSDUH survey.

NOTE: For each individual year, the MHSS or the NSDUH annual analysis weight was used to obtain the weighted estimates. For 2008-2012, the rescaled MHSS, the scaled NSDUH, or the unscaled NSDUH analysis weight for the combined 5-year data was used to obtain the weighted estimates. For 2008, MHSAMPWT was used to obtain the NSDUH estimates for the single year data as well as for the combined 5-year data.

[a] Difference between this estimate and the corresponding estimate from the 2008-2012 SCID data is statistically significant at the .05 level.

[1] Respondents with unknown past year MDE data were excluded.

[2] Among those with MDE. "Any Treatment" includes categories either "Saw or Talked to Health Professional" or "Prescription Medication."

[3] Mental health treatment/counseling is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.

[4] Respondents with unknown suicide information were excluded.

[5] Variables from this section were based on the SCID clinical interview, while other characteristic variables in Tables 6.1 to 6.10 were based on the NSDUH main interview.

[6] Three categories of the level of mental illness severity based on functional impairment are defined based on the SCID disorder diagnosis and GAF scales: mild mental illness, moderate mental illness, and serious mental illness. Any mental illness includes persons in any of the three categories. For more details, see Table 1.1 in Chapter 1.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table 6.6 Weighted Sample Description of SCID and CAI Respondents, by Substance Use Measures: 2008-2012

	2008		2009		2010		2011		2012		2008-2012		
	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	
Characteristic	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct. (unscaled)	Wt. Pct. (scaled)
Total (Numbers in Thousands)	224,923	224,923	227,207	227,207	229,273	229,273	232,625	232,625	235,124	235,124	231,890	229,830	231,890
Past Month Substance Use													
Illicit Drugs[1]	7.4	7.9	8.4	8.6	10.0	8.8	9.7	8.6	8.2	9.1	8.9	8.6	8.7
Marijuana	6.3	6.1	7.1	6.6	8.5	6.8	8.0	6.9	6.1	7.3	7.2	6.7	6.9
Illicit Drugs Other Than Marijuana[1]	2.4	3.3	5.3	3.6	3.2	3.5	2.4	3.0	2.9	3.4	2.8	3.3	3.3
Tobacco Products[2]	28.8	30.3	23.9	29.5	28.1	29.2	30.8	28.3	28.7	28.6	29.2	29.2	28.8
Alcohol	57.7	55.7	61.1	55.9	56.6	55.9	58.5	55.9	57.3	56.3	57.8	55.9	56.0
Past Year Substance Abuse or Dependence													
Alcohol or Drug Substance Use Disorder[3]	9.0	9.1	9.0	9.2	7.9	8.9	8.9	8.1	8.8	8.8	8.7	8.8	8.6
Alcohol Use Disorder	7.9	7.7	6.9	7.8	7.4	7.4	7.2	6.8	7.6	7.2	7.4	7.3	7.1
Illicit Drug Use Disorder	2.4	2.6	3.6	2.7	1.6	2.6	2.8	2.3	2.9	2.7	2.6	2.6	2.5
No Substance Use Disorder	91.0	90.9	91.0	90.8	92.1	91.1	91.1	91.9	91.2	91.2	91.3	91.2	91.4

CAI = computer-assisted interviewing; K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule; wt. pct. = weighted percent.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. The respondent characteristics in this table are based on data collected from the NSDUH CAI interview.

NOTE: For each individual year, the MHSS or the NSDUH annual analysis weight was used to obtain the weighted estimates. For 2008-2012, the rescaled MHSS, the scaled NSDUH, or the unscaled NSDUH analysis weight for the combined 5-year data was used to obtain the weighted estimates. For 2008, MHSAMPWT was used to obtain the NSDUH estimates for the single year data as well as for the combined 5-year data.

[a] Difference between this estimate and the corresponding estimate from the 2008-2012 SCID data is statistically significant at the .05 level.

[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other Than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically, including data from original methamphetamine questions but not including new methamphetamine items added in 2005 and 2006.

[2] Tobacco Products include cigarettes, smokeless tobacco (i.e., chewing tobacco or snuff), cigars, or pipe tobacco.

[3] Substance Use Disorder is defined as meeting criteria for illicit drug or alcohol dependence or abuse. Dependence or abuse is based on definitions found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table 6.7 Sample Description of SCID Respondents, by Past Year K6 Score Frequency Distribution: 2008-2012

	2008	2009	2010	2011	2012	2008-2012
Past Year K6 Score	**Unweighted Pct.**	**Unweighted Pct.**	**Unweighted Pct.**	**Unweighted Pct.**	**Unweighted Pct.**	**Unweighted Pct.**
0	2.4	4.0	10.7	11.4	11.1	8.2
1	1.1	1.9	6.2	7.6	6.4	4.9
2	1.7	1.9	6.8	7.0	7.3	5.2
3	1.3	3.5	4.5	6.6	5.5	4.4
4	3.6	4.8	6.8	5.4	5.6	5.1
5	3.1	2.9	3.5	5.6	4.7	4.3
6	3.8	5.0	4.3	3.9	3.8	4.0
7	3.4	5.2	4.5	3.9	3.8	3.9
8	6.7	8.1	3.1	3.3	4.1	4.8
9	7.4	4.4	2.9	3.9	3.0	4.5
10	7.7	5.6	1.2	2.5	4.0	4.5
11	6.1	6.9	3.9	4.6	3.0	4.7
12	11.2	9.8	3.3	5.5	5.0	7.1
13	7.4	7.5	6.4	3.3	3.8	5.2
14	5.3	3.8	3.9	3.5	3.9	4.1
15	5.6	6.3	3.7	3.1	3.7	4.3
16	4.1	3.7	4.1	2.7	3.6	3.6
17	2.9	1.7	3.7	3.1	3.6	3.1
18	5.0	4.0	5.2	3.3	4.7	4.4
19	1.5	2.1	2.1	2.4	2.1	2.0
20	1.8	1.5	2.1	1.4	2.1	1.8
21	1.3	1.5	1.2	1.0	0.6	1.0
22	0.9	1.2	0.8	1.2	1.4	1.1
23	0.9	0.2	1.6	0.8	0.5	0.8
24	3.8	2.3	3.9	2.7	2.8	3.1
Total (Sample Size)	1,500	520	516	1,495	1,622	5,653

CAI = computer-assisted interviewing; K6 = Kessler-6, a 6-item psychological distress scale; pct. = percent; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. Past year K6 total scores in this table are based on data collected from the NSDUH CAI interview.

NOTE: K6SCMAX was used in this table, which is defined as the higher K6 score between the past month K6 total score and the K6 total score in the worst month of the past year if the worst month was not the past 30 days.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 6.8 Weighted Sample Description of SCID and CAI Respondents, by Past Year K6 Score Frequency Distribution: 2008-2012

Past Year K6 Score	2008		2009		2010		2011		2012		2008-2012		
	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	
	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct. (unscaled)	Wt. Pct. (scaled)
0	15.2	22.0	20.7	22.9	25.1	23.3	22.7	23.9	20.2	23.5	21.2	23.1	23.4
1	9.8	10.8	3.8	10.3	11.2	10.5	12.3	10.9	9.9	10.4	10.7	10.6	10.6
2	9.9	10.7	7.9	10.9	9.7	10.4	10.4	11.1	12.7	10.7	11.0	10.8	10.8
3	9.4	9.5	15.5	8.7	9.4	9.1	11.2	8.7	8.2	8.9	9.8	9.0	8.9
4	8.8	7.9	5.9	7.8	6.8	7.9	6.2	7.6	9.8	7.6	7.8	7.8	7.7
5	7.3	6.1	6.6	6.1	5.1	5.8	7.2	5.7	5.4	6.2	6.2	6.0	5.9
6	6.7	5.3	4.5	5.5	5.1	5.2	3.9	5.4	3.6	5.4	4.3	5.4	5.4
7	6.9	4.1	11.4	4.0	6.3	3.9	3.9	3.7	5.2	3.7	5.3	3.9	3.8
8	4.2	3.3	3.1	3.5	2.0	3.4	2.4	3.2	4.5	3.2	3.3	3.3	3.2
9	4.4	2.7	0.7	2.6	3.2	2.9	2.5	2.7	2.9	2.6	2.9	2.7	2.7
10	2.2	2.6	3.2	2.4	0.9	2.4	1.8	2.4	2.2	2.4	1.9	2.4	2.4
11	1.6	2.0	1.5	2.3	2.5	2.1	2.4	2.1	1.8	2.1	2.1	2.1	2.1
12	3.8	2.7	5.9	2.8	1.2	2.8	2.7	2.5	2.8	2.6	2.8	2.7	2.6
13	2.2	1.6	1.7	1.6	2.0	1.5	1.3	1.5	1.8	1.6	1.7	1.6	1.6
14	0.8	1.4	0.8	1.4	1.2	1.3	1.5	1.3	1.5	1.3	1.3	1.3	1.3
15	1.4	1.1	1.7	1.2	0.9	1.2	1.4	1.0	1.0	1.1	1.2	1.1	1.1
16	1.0	1.1	0.9	1.1	1.8	1.1	1.2	1.1	1.4	1.0	1.3	1.1	1.1
17	0.8	1.0	0.4	0.9	1.1	0.8	1.1	0.9	0.8	0.9	0.9	0.9	0.9
18	1.3	1.4	1.3	1.2	1.2	1.5	1.2	1.2	1.5	1.5	1.4	1.4	1.4
19	0.4	0.5	0.6	0.6	0.4	0.6	0.7	0.6	0.6	0.7	0.6	0.6	0.6
20	0.6	0.5	0.7	0.4	1.2	0.5	0.4	0.5	0.6	0.6	0.6	0.5	0.5
21	0.5	0.4	0.4	0.4	0.2	0.3	0.3	0.3	0.1	0.4	0.2	0.4	0.3
22	0.1	0.3	0.2	0.3	0.2	0.3	0.3	0.3	0.5	0.3	0.3	0.3	0.3
23	0.1	0.2	0.0	0.2	0.3	0.2	0.2	0.2	0.3	0.3	0.2	0.2	0.3
24	0.8	0.8	0.7	0.9	1.0	1.0	0.8	1.1	0.8	1.0	0.8	1.0	1.0
Total (Numbers in Thousands)	224,923	224,923	227,207	227,207	229,273	229,273	232,625	232,625	235,124	235,124	231,890	229,830	231,890

CAI = computer-assisted interviewing; K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule; wt. pct. = weighted percent.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. Past year K6 total scores in this table are based on data collected from the NSDUH CAI interview.

NOTE: K6SCMAX was used in this table, which is defined as the higher K6 score between the past month K6 total score and the K6 total score in the worst month of the past year if the worst month was not the past 30 days.

NOTE: For each individual year, the MHSS or the NSDUH annual analysis weight was used to obtain the weighted estimates. For 2008-2012, the rescaled MHSS, the scaled NSDUH, or the unscaled NSDUH analysis weight for the combined 5-year data was used to obtain the weighted estimates. For 2008, MHSAMPWT was used to obtain the NSDUH estimates for the single year data as well as for the combined 5-year data.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table 6.9 Weighted Sample Description of SCID and CAI Respondents, by K6 Scores: 2008-2012

	2008		2009		2010		2011		2012		2008-2012		
	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	
K6 Score	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE) (unscaled)	Mean (SE) (scaled)
Past Month Total Score	4.11 (0.22)	3.77 (0.03)	4.27 (0.47)	3.79 (0.03)	3.53 (0.25)	3.79 (0.04)	3.66 (0.14)	3.70 (0.03)	3.92 (0.13)	3.83 (0.04)	3.81 (0.08)	3.77 (0.02)	3.77 (0.02)
Past Year Total Score	5.42 (0.30)	4.90 (0.04)	5.37 (0.50)	4.89 (0.04)	4.79 (0.29)	4.88 (0.04)	4.74 (0.17)	4.78 (0.04)	5.07 (0.14)	4.91 (0.04)	4.97 (0.10)	4.87 (0.02)	4.86 (0.02)

CAI = computer-assisted interviewing; K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; SE = standard error; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. Past month and Past year K6 total scores in this table are based on data collected from the NSDUH CAI interview.

NOTE: For each individual year, the MHSS or the NSDUH annual analysis weight was used to obtain the weighted estimates. For 2008-2012, the rescaled MHSS, the scaled NSDUH, or the unscaled NSDUH analysis weight for the combined 5-year data was used to obtain the weighted estimates. For 2008, MHSAMPWT was used to obtain the NSDUH estimates for the single year data as well as for the combined 5-year data.

[a] Difference between this estimate and the corresponding estimate from the 2008-2012 SCID data is statistically significant at the .05 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table 6.10 Weighted WHODAS Past Year Total Score Frequency Distribution of SCID and CAI Respondents, by Sampling Period: 2008-2012

WHODAS Total Score	2008		2009		2010		2011		2012		2008-2012		
	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	SCID	CAI	
	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct.	Wt. Pct. (unscaled)	Wt. Pct. (scaled)
0	47.4	46.5	38.8	45.6	49.7	46.0	46.0	47.1	43.4	45.8	45.5	46.2	46.4
1	5.8	7.0	2.3	7.2	7.6	7.1	7.3	6.9	6.3	7.3	6.6	7.1	7.1
2	3.5	6.9	11.6	7.1	7.8	7.0	9.0	7.2	9.3	6.6	8.4	7.0	6.9
3	11.3	5.7	10.6	5.9	5.4	5.8	5.2	5.7	5.7	5.6	6.3	5.7	5.7
4	7.9	4.7	4.8	5.1	4.0	4.8	4.0	4.7	3.9	4.9	4.4	4.9	4.8
5	3.0	4.3	8.8	4.1	4.3	4.4	4.0	4.2	5.6	4.1	4.7	4.2	4.2
6	4.1	4.0	2.1	3.8	2.8	3.6	3.4	3.5	3.6	3.7	3.4	3.7	3.7
7	2.9	3.0	6.2	3.2	1.6	3.0	3.4	2.8	3.3	3.3	3.2	3.1	3.0
8	1.7	3.3	2.8	3.3	2.8	3.4	3.9	3.5	3.4	3.3	3.3	3.4	3.4
9	1.1	2.4	0.9	2.2	2.1	2.3	2.0	2.2	1.4	2.1	1.6	2.2	2.2
10	1.2	2.1	2.1	1.9	1.0	2.1	2.1	1.8	2.7	2.2	2.0	2.0	2.0
11	0.8	1.6	1.3	1.7	1.6	1.7	1.4	1.5	2.0	1.6	1.6	1.6	1.6
12	1.9	1.3	0.9	1.6	1.3	1.3	1.7	1.4	1.5	1.5	1.6	1.4	1.4
13	1.9	1.2	1.1	1.2	1.4	1.3	1.3	1.2	1.3	1.2	1.4	1.2	1.2
14	1.8	0.9	1.1	1.2	2.1	1.1	0.8	1.1	1.0	1.1	1.2	1.1	1.1
15	0.8	0.9	0.8	0.9	0.8	0.9	1.0	1.0	1.2	1.0	1.0	1.0	1.0
16	0.7	1.2	0.9	1.0	0.4	0.9	1.2	1.0	1.0	1.2	0.9	1.1	1.1
17	0.2	0.6	0.3	0.5	0.1	0.7	0.4	0.7	0.6	0.7	0.4	0.6	0.7
18	0.4	0.4	0.3	0.6	0.4	0.5	0.6	0.6	1.1	0.7	0.7	0.6	0.6
19	0.3	0.4	0.3	0.4	0.6	0.4	0.3	0.4	0.6	0.4	0.4	0.4	0.4
20	0.4	0.3	0.1	0.3	0.6	0.4	0.2	0.3	0.3	0.3	0.3	0.3	0.3
21	0.2	0.3	1.5	0.3	0.3	0.3	0.4	0.3	0.3	0.3	0.4	0.3	0.3
22	0.5	0.3	0.2	0.3	0.3	0.3	0.2	0.2	0.2	0.3	0.2	0.3	0.3
23	0.2	0.2	0.2	0.2	0.8	0.2	0.1	0.3	0.3	0.2	0.3	0.2	0.2
24	0.2	0.3	0.0	0.3	0.4	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.4
Total Population (Numbers in Thousands)	224,923	224,923	227,207	227,207	229,273	229,273	232,625	232,625	235,124	235,124	231,890	229,830	231,890

CAI = computer-assisted interviewing; K6 = Kessler-6, a 6-item psychological distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th edition); SDS = Sheehan Disability Scale; WHODAS = 8-item World Health Organization Disability Assessment Schedule; wt. pct. = weighted percent.

NOTE: This table does not include any clinical interview (interviews were conducted using the SCID) respondent who had missing values for all K6 and WHODAS/SDS item scores in the NSDUH questionnaire. These respondents were not included in any of the model development analyses. WHODAS past year total scores in this table are based on data collected from the NSDUH CAI interview.

NOTE: For each individual year, the MHSS or the NSDUH annual analysis weight was used to obtain the weighted estimates. For 2008-2012, the rescaled MHSS, the scaled NSDUH, or the unscaled NSDUH analysis weight for the combined 5-year data was used to obtain the weighted estimates. For 2008, MHSAMPWT was used to obtain the NSDUH estimates for the single year data as well as for the combined 5-year data.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

7. Estimation of Mental Illness for the 2008 to 2012 NSDUHs

7.1 Background

The primary objective of the Mental Health Surveillance Study (MHSS) is to produce annual model-based estimates of serious mental illness (SMI) prevalence among the civilian, noninstitutionalized, adult U.S. population using data from the National Survey on Drug Use and Health (NSDUH). These estimates must be of sufficient accuracy, and the examination of trends over time should be possible. Secondary objectives include the estimation of other categories of mental illness, such as any mental illness (AMI), moderate mental illness (MMI), and low (mild) mental illness (LMI).

Although the ideal way to estimate SMI in NSDUH would be to administer a clinical diagnostic interview annually to all 45,000 adult respondents, this approach is not feasible because of constraints on the NSDUH interview time and the need for trained mental health clinicians to conduct the interviews. Therefore, the approach adopted by the Substance Abuse and Mental Health Services Administration (SAMHSA) was to administer a psychological diagnostic interview to a subsample of adult respondents who had completed the NSDUH interview as part of a clinical study from which model-derived estimates of SMI and AMI from the main NSDUH adult samples for each year from 2008 to 2011 were produced.

In 2008, the first year of the MHSS, approximately 1,500 NSDUH respondents participated in the clinical follow-up. The sample collected in 2008 was used to develop prediction models. Specifically, a weighted logistic regression model was fit on the clinical diagnostic data collected in 2008. The dependent variable for the model was whether or not the respondent had a diagnosis of SMI based on the clinical diagnostic interview. The predictor variables were mental health-related items collected in the main NSDUH interview. The model was used to produce a predicted probability of having SMI for each of the respondents to the clinical interview. A cut point was established among the predicted probabilities such that if adults with probabilities at or above the cut point were predicted to have SMI and the rest were not, the weighted number of false positives (adults not diagnosed to have SMI but predicted to have SMI) would come as close as possible to equaling the weighted number of false negatives (adults diagnosed to have SMI but not predicted to have SMI). Because the predictor variables in the model were variables collected on the NSDUH main interview, a probability of having SMI could be predicted for every NSDUH adult respondent using the estimated model parameters. Applying the cut point (determined from the clinical sample) on the predicted probabilities estimated in the NSDUH adult sample, each NSDUH adult respondent was classified as having or not having SMI. This dichotomous variable was then used to compute prevalence estimates of SMI for adults.

For comparability in estimating trends, the 2009, 2010, and 2011 SMI estimates were computed using the method described above. In addition, the estimated logistic model for SMI based on 2008 clinical data was used to determine a cut point for AMI (equalizing the weighted

AMI false positive and false negative as closely as possible in the 2008 clinical sample), which in turn was used to predict the AMI status for NSDUH main interview adult respondents from 2008 to 2011.

Although SAMHSA used the estimation methods based on the 2008 clinical follow-up data for producing annual estimates of SMI, the clinical data collection continued from 2009 to 2012. That is, MHSS clinical interview samples of roughly 500 cases in 2009 and 2010 were collected, followed by 1,500 cases in both 2011 and 2012. After 2012, the annual clinical interviews were discontinued. The data collected from 2008 to 2012 were used to determine whether the 2008 estimation methods should be revised; revisions to the methods would only be implemented if it could be shown that more accurate estimates of SMI could be produced.

Based on the analyses of the nearly 5,500 clinical interviews from 2008 through 2012, it was determined that changes to the 2008 estimation methods would produce considerably more accurate estimates of mental illness. As a result, the 2012 estimates were based on the revised estimation procedures. Also, previously published estimate of mental illness for 2008 through 2011 were revised. Although the clinical sample was discontinued after 2012, the 2012 estimation methods will be used to produce model-based estimates of mental illness annually.

This chapter provides a general overview of the estimation methodology used to produce the prevalence estimates of mental illness among adults aged 18 or older and describes the revisions to the estimation methodology that were implemented in 2012. Specifically, revisions to the estimation methodology included the following:

- the revision of the weights used in fitting the model in the clinical sample (the revisions to the model weights were described in detail in Chapter 5 of this report and are therefore not discussed in detail in this chapter);
- the addition of covariates into the logistic model for SMI to improve the accuracy of SMI predictions and to remove biases in prevalence estimates at the subpopulation level, particularly within age groups; and
- the use of the entire clinical sample from 2008 to 2012 in fitting the model for SMI.

Section 7.2 describes the 2008 estimation method. Section 7.3 provides an overview of the investigation into possible improvements of the 2008 model using the revised weights. A more detailed description of the methodological studies that led to the 2012 estimation methodology can be found elsewhere (Center for Behavioral Health Statistics and Quality [CBHSQ], in press a). Section 7.4 compares SMI and AMI prevalence estimates computed using the 2008 and 2012 estimation methods with their respective weighting schemes. Section 7.5 discusses some caveats of NSDUH model-based mental illness prevalence estimates.

7.2 2008 Estimation Method

As described previously, in 2008, a randomly selected subsample of approximately 1,500 adults who had completed the NSDUH main interview in English was recruited for a follow-up clinical interview consisting of a diagnostic assessment for mental disorders. In addition, as described in greater detail in the first three chapters, a split-sample design was incorporated into

the full 2008 NSDUH data collection in order to determine the optimal scale for measuring functional impairment in NSDUH. Roughly half of the adult respondents were assigned to receive an abbreviated eight-item version of the World Health Organization Disability Assessment Schedule (WHODAS) (Novak et al., 2010), and the other half were assigned to receive the Sheehan Disability Scale (SDS) (Leon et al., 1997). The half sample that received the WHODAS is referred to as the 2008A sample in this report, and the half sample that received the SDS is referred to as the 2008B sample.

Weighted logistic models that predicted SMI were developed for each half sample using the data from the subsample of clinical interview respondents. The short scales (the Kessler-6 [K6] in combination with the WHODAS for the 2008A subsample or the K6 in combination with the SDS for the 2008B subsample) were used as predictors in the weighted logistic models of mental illness assessed via the clinical interviews. The model parameter estimates then were used to predict SMI in the 2008 NSDUH full adult sample based on their responses to the short scale questions.

Based on an analysis of the 2008 MHSS data, it was determined that the WHODAS was the better predictor of SMI and that this scale would be used in combination with the K6 scale to predict SMI. It also was decided that the WHODAS would continue to be administered as the sole impairment scale in the 2009 and subsequent NSDUHs (Aldworth et al., 2009). Therefore, the remainder of this section focuses on the model that was fit on the half sample that was assigned to the WHODAS. This model is called the "2008 (WHODAS) model." A similar model was constructed in the 2008 MHSS for the half sample that was assigned to the SDS. See Section 7.3.5 for more details on the SDS half sample.

Originally, a weighted logistic regression model was fit on clinical diagnostic data collected from the 2008A subsample (n = 759). The dependent variable for the model was a clinical diagnosis of SMI (1 = has SMI, 0 = does not have SMI), and the predictor variables were alternative scores based on the psychological distress (K6) and function impairment (WHODAS) items collected in the NSDUH main interview. The model was used to produce a predicted probability of having SMI for each clinical interview respondent. A cut point was then established among the predicted probabilities of having SMI such that if adults with probabilities at or above the cut point were predicted to have SMI and the rest not, the weighted number of adults in the MHSS clinical sample incorrectly predicted to have SMI (false positives) would come as close as possible to equaling the weighted number of adults incorrectly predicted to not have SMI (false negatives). This approach ensures that the direct estimate of SMI based on the clinical diagnoses would be as close as possible to the predicted estimate based on the predicted SMI status for the clinical interview respondents.

Because the predictor variables in the model were variables collected during the main NSDUH interview, a probability of having SMI can be computed for every NSDUH adult respondent using the estimated model parameters. Employing the predicted probabilities from the adult NSDUH main study sample, SMI predicted values were computed for adult respondents in NSDUH main interview samples. That is, predicted values of SMI were produced (1 = predicted to have SMI; 0 = predicted not to have SMI) by applying the cut point to the individual respondent's predicted probabilities. The SMI predicted values for NSDUH

respondents then were used to compute prevalence estimates of SMI for adults. The resulting prevalence estimator is referred to as the *cut point estimator* in the remainder of this report.

The cut point determined from fitting SMI in the 2008 WHODAS clinical sample was applied to respondents in NSDUH main interview samples from 2009 and beyond to estimate SMI prevalences annually. Moreover, the probabilities of having SMI estimated from the 2008 model were used to make AMI predictions for adult NSDUH respondents. This was done by determining a second cut point such that if adults with probabilities at or above the cut point were predicted to have AMI and the rest not, the weighted total of false AMI positives and false AMI negatives in the clinical sample would come as close as possible to being equal.

Finally, a third cut point was determined so that if adults with probabilities at or above the cut point were predicted to have either serious or moderate mental illness (SMMI) and the rest not, the weighted total of false SMMI positives and false SMMI negatives in the clinical sample would come as close as possible to being equal. Predicted values for adult NSDUH respondents could then be computed for moderate mental illness (MMI = SMMI but not SMI), mild mental illness (LMI = AMI but not SMMI), and no mental illness (not AMI).

7.3 2012 Estimation Method

7.3.1 Investigating the 2008 (WHODAS) Model and Estimation Methods

By the end of 2012, a nationally representative sample of approximately 5,000 completed clinical assessments had been collected that included the 759 respondents from 2008, 520 from 2009, 516 from 2010, 1,495 from 2011, and 1,622 from 2012. This larger dataset was used to evaluate the accuracy of SMI estimates based on the 2008 estimation methods. Based on this evaluation, it was determined that revised weights and a revised model would be used to compute more accurate mental illness prevalence estimates for 2012 (see Chapter 5 for a treatment of the changes in the weights). In addition, the previously released (2008 to 2011) estimates of mental illness would be revised using the improved methods. This would not only allow for the assessment of trends in mental illness prevalence across survey years, but also provide better annual measures of mental illness across all adults and within particular subpopulations.

7.3.2 Choosing Covariates in a Model for SMI

A variety of variables was considered as predictors in a revised model of SMI used to produce cut point estimators for SMI and AMI. The following criteria were used to decide on the number and type of variables that could be included in the model. First, only a limited number of covariates could reasonably be added to a logistic model based on clinical sample data having at most, 100 effective degrees of freedom (*df*) after collapsing strata in the NSDUH full adult sample (see Section 3.7 for more information) to ensure that no primary sampling unit (PSU) is

empty (100 variance strata with two variance PSUs each). A maximum of 10 (i.e., ≈ √ 100) variables were considered for inclusion in each model.[44]

Second, some variables closely related to mental illness were avoided because including them in the model could bias the cut point estimator for a domain (i.e., subpopulation) of interest. For example, although having received services for mental health is correlated with SMI, using mental health service receipt as a covariate in a model for SMI would produce a cut point that overpredicts SMI prevalence among adults having received mental health services. Moreover, including it in the model for SMI would prevent researchers from accurately measuring changes in the receipt of treatment among those with mental illness over time.

Ultimately, the goal was to construct a parsimonious model that could be used annually to assess changes in SMI (and other categories of mental illness) within and between sociodemographic groups. A covariate related to a domain of interest would tend to fix the relationship between SMI and that domain.[45] For example, if being employed resulted in an estimated 2 percent decrease in the odds of having SMI when all other things were equal, then treating that 2 percent decrease as fixed over time would impede the measurement of any changes in the relationship between SMI and employment. Therefore, in most instances variables closely related to domain-membership indicators were not considered for inclusion into the model as covariates.

7.3.3 Metrics Used to Evaluate Models of SMI: Error Rate and Bias Measure

With the above criteria in mind, the SMI models based on the different predictor variables were evaluated using mainly two metrics: the overall error rate and domain-level bias. The error rate is a measure of the predictive power of a cut point estimator based on a particular model. It is the sum of the estimated fraction of false positives and false negatives in the adult population. Models with lower error rates produce more accurate predictions of SMI than models with higher error rates.

Different combinations of the K6 and WHODAS items and scores were evaluated as predictors in a variety of models for SMI, but none led to meaningful reductions in the error rate when compared with the alternative K6 and WHODAS scores used in the 2008 model. By contrast, the addition of two variables from the main NSDUH interview, serious thoughts of suicide in the past year and the experience of a major depressive episode (MDE) in the past year, did noticeably decrease the error rate and were therefore to be included in a (potential) new model from SMI.

---

[44] This derives from the notion that the number of PSUs minus that number of strata must be greater than the number of estimated coefficients for the asymptotic properties of modeling fitting to be relevant. If the latter is p, the former should be at least p^2.

[45] The domains of interest were gender, age group, race/ethnicity group, region, county type, employment level, education, whether the adult received mental health treatment, whether the adult had health insurance, and the adult's household income in relation to the poverty threshold. Definitions of these domains can be found later in the text and in Table 7.3's footnotes.

The second metric used to evaluate the model was a measure of bias. A cut point estimator for SMI prevalence is based on a model, and a model is only as good as the assumptions on which it is based. As a result, a cut point estimator, unlike a model-free direct estimator computed directly from the clinical sample, can be systematically biased. A detailed description of other model-based estimators for SMI and AMI is contained in CBHSQ (in press a).

The bias measure for an SMI prevalence estimate can be measured among all adults or for domains. The bias measure for a domain was defined as the difference between the weighted proportions across the clinical sample within the domain of respondents predicted to have SMI and those actually diagnosed to have SMI (this is equal to the difference between the false positive rate and the false negative rate in the domain). Under the null hypothesis that there is no bias in the domain, this bias measure would not be significantly different from zero.

Mathematically, let y_k, c_k, w_k, and ω_k be the actual SMI diagnosis (1 = yes, 0 = no), the model-based SMI prediction, the main NSDUH weight, and the clinical sample weight, respectively, for respondent k. The cut point estimator for a domain-level prevalence is $C_D = \sum_D w_k c_k / \sum_D w_k$, where $\sum_D$ denotes summation over the domain. C_D has the same asymptotic expectation as $C_D' = \sum_D \omega_k c_k / \sum_D \omega_k$. The bias measure for C_D is B_D, the difference between C_D' and the asymptotically unbiased direct estimator $\sum_D \omega_k y_k / \sum_D \omega_k$. Note that $B_D = \sum_D \omega_k (c_k - y_k) / \sum_D \omega$ can be viewed as a simple weighted mean of $c_k - y_k$. The standard error (SE) of this mean can then be measured assuming each $c_k - y_k$ is an independent random variable with a common mean and an unknown variance that can vary across the k. Also, the ratio of B_D and its SE is asymptotically standard normal under the null hypothesis of no bias at the domain level. This ratio (and the normality assumption) was used as the t statistic in testing for the bias of a cut point estimator at the domain level.

As discussed, the cut point for SMI using the 2008 estimation method was determined so that the estimated proportion of false positives (adults predicted to have SMI but did not) and false negatives (adults predicted not to have SMI but did) in the clinical sample were as close to equal as possible. This property removed the possibility of systematic bias in the estimated proportion of adults having SMI in 2008 based on the cut point estimator of NSDUH respondents. Unfortunately, using the 2008 SMI cut point among all adults did not ensure the near equality of estimated false positives and false negatives among all adults for years other than 2008 or among domains for which SMI estimates are computed such as age groups. As a result, it was possible for the cut point estimator for a domain to be biased.

Bias in estimates of SMI was investigated for a number of domains of interest (i.e., gender, age group, race/ethnicity group, region, county type, employment level, education, whether the adult received mental health treatment, whether the adult had health insurance, and the adult's household income in relation to the poverty threshold). To do this, a model of SMI was fit on the entire WHODAS clinical sample from 2008 to 2012. Predictor variables in this model of SMI included past year K6, WHODAS, MDE, and suicidal thoughts. Results of the investigation indicated that SMI estimates within certain age groups were biased. The bias in the SMI estimates by these age groups is displayed in Table 7.1. As well as bias measures, Table 7.1

displays direct estimates of SMI[46] computed directly from the MHSS clinical sample using the clinical diagnoses and cut point estimates derived from the main NSDUH adult sample.

Table 7.1 Cut Point Estimates of SMI for Age Groups and Their Bias Measures Using a Model of SMI with Past Year K6, WHODAS, MDE, and Suicidal Thoughts as Predictor Variables: 2008A-2012 MHSS

Age	Direct Estimate[1]	Cut Point Estimate[2]	Bias Measure[3]	SE of Bias Measure[4]	*t* Value of Bias Measure[5]	*P* Value of Bias Measure[6]
All Adults	3.93	3.90	0.01	0.28	0.25	0.962
18 to 25	3.77	5.79	2.25	0.81	2.77	0.006
26 to 34	4.35	4.70	0.42	0.65	0.64	0.521
35 to 49	5.74	4.62	-1.56	0.58	-2.71	0.007
50 or Older	2.74	2.51	0.03	0.40	0.08	0.936

K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

[1] Direct estimate is computed using clinical diagnoses from the clinical respondent subsample only using the weights designed for direct estimation of this sample (MHFAAWGT; see Chapter 5).

[2] Cut point estimate is computed from the main NSDUH sample of adults using predicted SMI statuses with a scaled version of the main NSDUH weights to make it comparable with the direct estimate. The scaling factors were .12 for the 2008A sample, .04 for the 2009 sample, .14 for the 2010 sample, and .35 for the 2011 and 2012 samples.

[3] Bias measure is the weighted mean value of the difference between the true value of SMI and the predicted value of SMI taken across the clinical respondent subsample within the age group (see the main text of this chapter).

[4] SE is measured as discussed in the main text of this chapter.

[5] The *t* value is the bias measure divided by its SE.

[6] The *p* value is for a two-sided test of whether the bias measure is significantly different from zero.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; NSDUH main study and clinical sample, 2008-2012.

As demonstrated in Table 7.1, the direct estimate is slightly larger than the cut point estimate for all adults (3.93 vs. 3.90 percent, respectively). The cut point estimate for the 18 to 25 age group, however, is much higher (5.79 percent) than the direct estimate (3.77 percent) for this age group. Moreover, the measure of the bias of this estimate is significantly different from zero at the .01 level ($p = 0.006$). These bias results, which were not paralleled in other domains of interest (see Chapter 4 of CBHSQ, in press a), suggested that adding an age-related predictor variable or variables to the model for SMI was needed to remove the bias in SMI estimates within these age groups.

It is not very plausible that, when all other factors held constant, the probability of having SMI would change suddenly when an adult aged a single year (i.e., changed from one age group to another). Consequently, a number of continuous age variables (e.g., AGE1830) were considered for addition to the SMI model. Each of these candidate variables was then compared in terms of its impact on the age group bias measures and the overall error rates for SMI, AMI, and SMMI.

The age variable that was ultimately added to the SMI model (termed AGE1830) is a recoded version of a continuous age variable for adults. It is coded as either 12 or the difference between the respondent's age and 18, whichever was smaller. The variable increased as the

[46] Mathematically, $Y_D = \sum_D \omega_k y_k / \sum_D \omega_k$.

respondent aged from 18 to 30, but then leveled off at 12 after age 30. Adding this age variable led to a cut point that both equalized false positives and false negatives for all adults and roughly within adult age groups.

After evaluating a wide variety of model specifications for both SMI and AMI that focused on finding a model that had minimum values of bias measures and error rates, the final model chosen included the following predictor variables: past year K6 score, past year WHODAS score, age, past year MDE, and past year suicidal thoughts. Table 7.2 shows how adding variables to the original 2008 WHODAS model specification, which only included the K6 and WHODAS, increased the accuracy of the SMI prevalence estimate by reducing the error rate. The estimates in this table for all adults were computed using the WHODAS sample from 2008 through 2012. The final model displayed is the 2012 (WHODAS) model.

Table 7.2 SMI Cut Point Estimates for All Adults under Different Models: 2008 to 2012 MHSS

Model Specification	Cut Point Estimate	Error Rate	False Positive Rate	False Negative Rate
K6 + WHODAS (Refit 2008 WHODAS Model)	3.97	4.42	2.21	2.21
K6 + WHODAS + AGE1830	3.89	4.30	2.17	2.12
K6 + WHODAS + MDE + SUICIDAL THOUGHTS	3.90	3.97	1.99	1.98
K6 + WHODAS + AGE1830 + MDE + SUICIDAL THOUGHTS (2012 WHODAS Model)	3.95	3.84	1.92	1.93

Age1830 = difference between respondent's age and 18 or 12, whichever was smaller; K6 = Kessler-6, a 6-item psychological distress scale; MDE = past year major depressive episode; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness; SUICIDAL THOUGHTS = suicidal thoughts in the past year.

NOTE: Scaled weights were used for both the main NSDUH sample and the clinical sample.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; NSDUH main study and clinical sample, 2008-2012.

7.3.4 2012 WHODAS Model

The 2012 (WHODAS) model for SMI was fit with data from 4,912 clinical interview respondents who received WHODAS in the NSDUH main interview from 2008 through 2012 with the weights designed for analyzing the entire 2008-2012 WHODAS clinical sample (MHFAAWGT, see Section 5.5 in Chapter 5). The response variable Y equaled 1 when an SMI diagnosis was positive based on the clinical interview; otherwise, Y was 0. Letting $\mathbf{X}$ be a vector of characteristics attached to a NSDUH respondent and letting the probability that this respondent had SMI be $\pi = \Pr(Y = 1 \mid \mathbf{X})$, the 2012 model was

$$\text{logit}(\hat{\pi}) = \log[\hat{\pi} / (1 - \hat{\pi})] = -5.972664 + 0.0873416X_k + 0.3385193X_w + 1.9552664X_s + 1.1267330\,X_m + 0.1059137X_a \quad (7.1)$$

or

$$\hat{\pi} = \frac{1}{1 + \exp[-(-5.972664 + 0.0873416X_k + 0.3385193X_w + 1.9552664X_s + 1.1267330\,X_m + 0.1059137X_a)}},$$

where $\hat{\pi}$ is the estimated probability an adult had SMI. The covariates in equation (7.1) come from the main NSDUH interview data:

- X_k = *Alternative Past Year K6 Score* (WSPDSC2[47]): Past year K6 score of less than 8 recoded as 0; past year K6 score of 8 to 24 recoded as 1 to 17.
- X_w = *Alternative WHODAS Score* (WHODASC3): WHODAS item score of less than 2 recoded as 0; WHODAS item score of 2 to 3 recoded as 1, then summed for a score ranging from 0 to 8.
- X_s = *Serious Thoughts of Suicide in the Past Year* (MHSUTK_U): Coded as 1 if "yes"; coded as 0 otherwise.
- X_m = *Past Year MDE* (AMDEY2_U): Coded as 1 if the criteria for past year MDE were met; coded as 0 otherwise.
- X_a = *AGE1830* (Recoded Age): Coded as age minus 18 if aged 18 to 30; coded as 12 otherwise.

A cut point probability π_0 was determined, so that if $\hat{\pi} \geq \pi_0$ for a particular respondent, then he or she was predicted to be SMI positive; otherwise, he or she was predicted to be SMI negative. The cut point (0.260573529) was chosen so that the weighted numbers of false positives and false negatives in the MHSS dataset were as close to equal as possible. The predicted SMI status for all adult NSDUH respondents was used to compute prevalence estimates of SMI.

A second cut point probability (0.0192519810) was determined so that any respondent with an SMI probability greater than or equal to the cut point was predicted to be positive for AMI, and the remainder were predicted to be negative for AMI. The second cut point was chosen so that the weighted numbers of AMI false positives and false negatives were as close to equal as possible.

Estimates of SMMI (GAF score below 60) were analogously computed with the SMI method; the cut point was 0.077686285365. Estimates of LMI and MMI were derived by a process of subtraction. Respondents were classified as belonging to the MMI category if they belonged to the SMMI category, but they did not belong to the SMI category. Respondents were classified as belonging to the LMI category if they belonged to the AMI category, but not to the SMMI category.

7.3.5 2012 Model for the SDS Sample

One of the predictors in the original 2008 SDS model for SMI was a variable based on the SDS, similar to the WHODAS variable in the WHODAS-based model. This model was selected so that it provided SMI estimates that were comparable with the 2008 WHODAS model-based SMI estimates. However, when that model was applied to the 2008B MHSS clinical sample to produce an AMI prevalence estimate for all adults, the result (16.73 percent) was not close enough to the AMI estimates for 2008A and 2009 calculated using the 2008

[47] See Section 4.4 in Chapter 4 for information on how the missing values were treated for the predictor variables in the 2012 model.

WHODAS model (19.48 and 19.86 percent, respectively). See Aldworth et al. (2010, pp. 36-38). As a result, AMI (and SMMI) prevalence estimates derived from the 2008B sample were not previously reported, for example, in the 2011 NSDUH mental health findings report (CBHSQ, 2012c).

This led to a search for a better model for the 2008 SDS sample so that the data from this sample could be used when producing estimates of SMI and AMI in 2008. The resulting 2012 model for SDS had the same covariates as the 2012 WHODAS model except that the WHODAS variable was removed. In addition, for the 2008B respondents, an adjusted version of MDE variable was used.[48] The model was fit to the entire clinical sample from 2008 through 2012 with the weights designed for analyzing the entire 2008-2012 clinical sample (MHFNLWGT, see Section 5.5 in Chapter 5). This model was then used to predict mental illness statuses for the 2008B main NSDUH sample.

In order to determine whether the predicted values from the 2008B sample could be combined with predicted values from the complete WHODAS sample when making mental illness prevalence estimates, an investigation was conducted to determine whether the estimates produced using this 2012 SDS model differed from those using the 2012 WHODAS model. Also investigated was whether the estimates produced using this 2012 SDS model significantly differed from direct estimates. Tables 7.3 and 7.4 show small differences between domain and all-adult estimates for 2008 SMI and AMI derived from the 2008A NSDUH respondent sample using the 2012 WHODAS model and the 2008B NSDUH respondent sample using the 2012 SDS model. The tables also contain the statistics on the bias measure for the 2008B sample estimates. No bias measure is significantly different from zero at the 0.1 level. Consequently, it was determined that the predicted values from the 2008B sample could be combined with predicted values from the complete WHODAS sample for 2008A and for 2009 through 2012 when making mental illness prevalence estimates.

The 2012 SMI prediction model for the SDS sample is

$$\text{logit}(\hat{\pi}) = \log[\hat{\pi} / (1 - \hat{\pi})] = -5.7736246 + 0.1772067X_k + 1.8392433X_s + 1.6428623X_m + 0.1231266X_a \qquad (7.2)$$

or

$$\hat{\pi} = \frac{1}{1 + \exp[-(-5.7736246 + 0.1772067X_k + 1.8392433X_s + 1.6428623X_m + 0.1231266X_a)}.$$

All of the covariates in equation (7.2) appeared in equation (7.1) as well.

[48] The variable was developed to remove the context effects on MDE in the 2008B NSDUH instrument. The MDE data for respondents in the SDS sample were not comparable with the MDE data for the respondents in the WHODAS sample. This was because respondents differentially answered the MDE questions (i.e., context effects) based on whether they had answered the WHODAS or SDS item in the preceding questionnaire module. For more details, see Aldworth et al. (2012).

Table 7.3 Comparing Cut Point Estimates for SMI from the 2008B Sample Using the 2012 Model for the SDS Sample with Estimates from the 2008A Sample Using the 2012 Model for the WHODAS Sample

	Cut Point Estimates			Bias Measures for 2008B Estimates		
Domain	2008A	2008B	Difference	Bias Measure	Standard Error of Bias Measure	*P* Value of Bias Measure
All Adults	3.72	3.70	-0.02	-1.43	1.91	0.455
Age: 18 to 25	4.15	3.42	-0.73	-0.77	0.83	0.355
Age: 26 to 34	4.57	4.83	0.26	-0.48	2.16	0.825
Age: 35 to 49	4.74	4.92	0.18	2.11	2.74	0.442
Age: 50 or Older	2.52	2.52	0.00	-4.37	4.17	0.294
Health Insurance: Yes	3.54	3.44	-0.10	-0.68	2.02	0.736
Health Insurance: No	4.71	5.17	0.46	-9.14	5.79	0.115
< 100% of the Poverty Threshold[1]	6.25	5.61	-0.64	-16.80	11.29	0.136
100%-199% of the Poverty Threshold[1]	4.84	5.11	0.27	0.15	1.31	0.912
≥ 200% of the Poverty Threshold[1]	3.02	3.02	0.00	1.21	1.16	0.296
White, Not Hispanic	4.14	4.27	0.13	-2.46	2.47	0.319
Black, Not Hispanic	2.66	2.31	-0.35	-1.01	4.80	0.833
Other, Not Hispanic	3.40	2.46	-0.94	5.19	10.18	0.610
Hispanic	2.58	2.56	-0.02	0.31	0.38	0.418
Male	2.76	2.50	-0.26	-0.34	1.74	0.846
Female	4.61	4.83	0.22	-2.45	3.29	0.457
Northeast	3.36	4.35	0.99	-4.96	5.49	0.367
Midwest	4.23	3.75	-0.48	-0.02	0.56	0.970
South	3.71	3.14	-0.57	-2.46	2.09	0.239
West	3.53	4.02	0.49	4.82	4.41	0.275
Large Metro	3.55	3.60	0.05	1.71	1.71	0.316
Small Metro	3.82	4.15	0.33	-3.77	4.60	0.413
Nonmetro	4.07	3.19	-0.88	-3.84	2.86	0.179
Received Mental Health Treatment: Yes[2]	18.39	17.90	-0.49	-16.70	11.23	0.137
Received Mental Health Treatment: No[2]	1.43	1.50	0.07	0.94	0.91	0.301
Employed Full Time	2.69	3.09	0.40	0.95	1.68	0.571
Employed Part Time	4.06	3.40	-0.66	0.37	1.09	0.735
Unemployed	7.98	5.70	-2.28	0.98	2.31	0.672
Other Employment Status[3]	5.00	4.70	-0.30	-6.01	5.07	0.235
Less than High School	3.28	3.47	0.19	-10.30	7.07	0.144
High School Graduate	4.38	3.46	-0.92	0.25	0.53	0.635
Some College	4.33	4.88	0.55	-3.78	5.52	0.494
College Graduate	2.65	3.04	0.39	1.53	2.61	0.558

NSDUH = National Survey on Drug Use and Health; SDS = Sheehan Disability Scale; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

[1] Poverty Threshold incorporates information on family income, size, and composition and is calculated as a percentage of the U.S. Census Bureau's poverty thresholds. Respondents aged 18 to 22 who were living in a college dormitory were excluded.

[2] Received Mental Health Treatment is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.

[3] Other Employment includes students, persons keeping house or caring for children full time, retired or disabled persons, or other persons not in the labor force.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; NSDUH main study and clinical sample, 2008-2012.

Table 7.4 Comparing Cut Point Estimates for AMI from 2008B Sample Using the 2012 Model for the SDS Sample with Estimates from the 2008A Sample Using the 2012 Model for the WHODAS Sample

	Cut Point Estimates			Bias Measures for 2008B Estimates		
Domain	2008A	2008B	Difference	Bias Measure	Standard Error of Bias Measure	*P* Value of Bias Measure
All Adults	17.69	17.78	0.09	-1.58	2.75	0.566
Age: 18 to 25	18.83	18.44	-0.39	-2.17	3.48	0.532
Age: 26 to 34	20.75	23.14	2.39	-11.20	7.77	0.150
Age: 35 to 49	20.08	20.19	0.11	4.17	5.02	0.405
Age: 50 or Older	14.43	13.79	-0.64	-0.93	4.57	0.839
Health Insurance: Yes	17.00	16.62	-0.38	-1.18	2.97	0.692
Health Insurance: No	21.43	24.44	3.01	-5.76	5.67	0.310
< 100% of the Poverty Threshold[1]	25.01	25.72	0.71	-18.90	11.51	0.100
100%-199% of the Poverty Threshold[1]	22.39	21.15	-1.24	6.02	4.43	0.174
≥ 200% of the Poverty Threshold[1]	15.27	15.60	0.33	0.07	2.87	0.981
White, Not Hispanic	18.44	18.10	-0.34	-3.16	3.44	0.359
Black, Not Hispanic	15.79	17.60	1.81	-10.60	6.56	0.105
Other, Not Hispanic	16.70	16.89	0.19	17.86	9.85	0.070
Hispanic	15.91	16.70	0.79	4.75	7.42	0.522
Male	13.45	13.80	0.35	0.84	3.84	0.827
Female	21.64	21.50	-0.14	-3.84	3.90	0.324
Northeast	16.68	21.16	4.48	-6.49	7.11	0.361
Midwest	18.64	18.04	-0.60	4.97	2.71	0.067
South	17.50	16.49	-1.01	-6.67	3.82	0.081
West	17.89	16.86	-1.03	7.35	7.28	0.313
Large Metro	16.91	17.57	0.66	-1.09	4.00	0.785
Small Metro	18.14	18.61	0.47	-3.72	5.69	0.514
Nonmetro	19.35	16.93	-2.42	1.00	2.99	0.738
Received Mental Health Treatment: Yes[2]	53.92	53.85	-0.07	-2.19	14.34	0.879
Received Mental Health Treatment: No[2]	12.05	12.16	0.11	-1.49	2.26	0.511
Employed Full Time	15.01	16.38	1.37	-0.97	3.83	0.800
Employed Part Time	20.86	16.96	-3.90	0.33	4.15	0.936
Unemployed	23.72	25.45	1.73	0.74	13.94	0.958
Other Employment Status[3]	20.55	19.66	-0.89	-3.49	5.48	0.524
Less than High School	19.09	21.57	2.48	-0.68	6.17	0.912
High School Graduate	17.77	16.31	-1.46	-0.11	2.43	0.964
Some College	18.55	20.13	1.58	-6.82	7.10	0.337
College Graduate	16.04	15.13	-0.91	2.07	5.33	0.698

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health; SDS = Sheehan Disability Scale; WHODAS = World Health Organization Disability Assessment Schedule.

[1] Poverty Threshold incorporates information on family income, size, and composition and is calculated as a percentage of the U.S. Census Bureau's poverty thresholds. Respondents aged 18 to 22 who were living in a college dormitory were excluded.

[2] Received Mental Health Treatment is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.

[3] Other Employment includes students, persons keeping house or caring for children full time, retired or disabled persons, or other persons not in the labor force.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; NSDUH main study and clinical sample, 2008-2012.

7.4 Comparisons of SMI and AMI Estimates Using the 2012 and 2008 Estimation Methods

Chapter 5 and the previous sections of this chapter discuss why and how the methods of weighting the clinical sample data and then modeling that data to produce SMI and AMI estimates have changed. This section investigates the effects of changes in estimation methodology on the SMI and AMI estimates. Tables 7.5 and 7.6 display SMI and AMI prevalence estimates for the combined time period 2008-2012 computed using the opposing estimation methods, that is, the NSDUH adults mental illness estimates using the 2008 model and weights versus the 2012 model and weights.

The SMI estimates for all adults in the NSDUH sample in these two tables for 2008-2012 have been computed using NSDUH analysis weights. In this combined prevalence estimate, the estimate from 2012 "counted" slightly more than the estimate from 2008 because the adult population size in 2012 was slightly larger. The estimates from the 2008 WHODAS sample (the 2008A sample) and SDS (2008B) sample were each implicitly weighted by half the number of adults in 2008 with one exception. Only the 2008A sample was used for AMI estimates with the 2008 method, with the estimates of that half sample then implicitly scaled by the adult population in 2008.

By contrast, the direct estimates in the tables were computed with the final scaled clinical sample weights as described in Section 5.5. The final clinical sample weights were scaled for use in analyses with combined years of data to increase statistical efficiency.[49] This scaling has the potential of biasing the resulting estimates. An *F* test applied to the all-adult estimates based on the clinical sample revealed no significant differences among the years in either SMI or AMI prevalences at the 0.1 level.

Table 7.5 presents the estimates of SMI for the overall population and within several domains of interest. As shown in the table, the 2012-method-based SMI estimates for the adult population and within domains tended to be smaller than their corresponding 2008-method-based counterparts. This is due mostly to the changes in the weights used in fitting the models. Recall that there should be almost no bias in the standard cut point estimates for all adults due to model fitting because of attempts to equalize the weighted false positive and false negative rates. Nevertheless, the 2008 method-based estimate has a significant upward bias measure, which must be the result of it being based on a fit using different weights.

Most of the 2012 method-based estimates in Table 7.5 were closer to the nearly unbiased direct estimates from the clinical sample than the 2008-method-based estimates. This is especially true for the 18 to 25 age group where the estimate using the 2008 method has a large bias. Only one domain in Table 7.5, the poverty level of less than a 100 percent threshold, had a significantly biased SMI estimate when applying the 2012 method. That is not wholly unexpected under the null hypothesis of no domain-level biases, given the number of tests conducted and therefore type I error inflation (64).

[49] The scaling factors were .06 for the 2008A and 2008B clinical samples, .04 for the 2009 sample, .14 for the 2010 sample, and .35 for the 2011 and 2012 samples.

Table 7.5 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008-2012 Adult NSDUH Main Study and Clinical Sample

Domain	Direct Estimate from Clinical Sample	Estimate Using 2008 Models	Estimate Using 2012 Model
Overall	3.9	4.9[a]	3.9
Age			
18 to 25	3.7	7.7[a]	3.8
26 to 34	4.2	6.5[a]	5.0
35 to 49	5.7	5.5	5.0
50 or Older	2.8	3.0	2.8
Gender			
Male	3.0	3.4	2.9
Female	4.9	6.3[a]	4.8
Race			
White, Not Hispanic	4.4	5.2	4.3
Black, Not Hispanic	3.5	3.9	3.1
Other, Not Hispanic	4.5	4.4	3.1
Hispanic	2.0	4.4	3.1
Region			
Northeast	3.1	4.5	3.7
Midwest	4.0	5.1[a]	4.2
South	3.9	4.8	3.7
West	4.8	5.2	4.1
County Type			
Large Metro	3.7	4.6	3.6
Small Metro	4.2	5.2[a]	4.2
Nonmetro	4.2	5.3	4.2
Received Mental Health Treatment[1]			
Yes	19.4	21.3	18.5
No	1.5	2.3[a]	1.6
Employment Level			
Employed Full Time	2.4	3.6[a]	2.8
Employed Part Time	4.2	5.5[a]	4.0
Unemployed	5.3	7.8	6.3
Other[2]	6.2	6.3	5.2
Education			
Less than High School	5.8	5.5	4.0
High School Graduate	3.8	5.1	4.0
Some College	4.5	5.9[a]	4.5
College Graduate	2.9	3.5	3.1

See notes at end of table. (continued)

Table 7.5 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008-2012 Adult NSDUH Main Study and Clinical Sample (continued)

Domain	Direct Estimate from Clinical Sample	Estimate Using 2008 Models	Estimate Using 2012 Model
Poverty Threshold[3]			
< 100% Threshold	9.9	9.1	6.9[a]
100% - 199% Threshold	5.3	6.1[a]	4.8
≥ 200% Threshold	2.5	3.7[a]	3.0
Health Insurance			
Yes	3.6	4.6[a]	3.6
No	5.8	6.7	5.2

NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness.

NOTE: The weights were *not* scaled for the NSDUH estimators when combining data across years. For the direct estimator, the scaling factors were .06 for the 2008A and 2008B clinical samples, .04 for the 2009 sample, .14 for the 2010 sample, and .35 for the 2011 and 2012 samples.

[a] The difference between the model-based and direct estimates computed from the subclinical sample is significantly different from zero at the .05 level.

[1] Received Mental Health Treatment is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.

[2] Other Employment includes students, persons keeping house or caring for children full time, retired or disabled persons, or other persons not in the labor force.

[3] Poverty Threshold incorporates information on family income, size, and composition and is calculated as a percentage of the U.S. Census Bureau's poverty thresholds. Respondents aged 18 to 22 who were living in a college dormitory were excluded.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; NSDUH main study and clinical sample, 2008-2012.

Table 7.6 presents the estimates of AMI for the overall population and within the same domains of interest (AMI estimates computed with the 2008 model do not use data from the 2008B sample). Similar to the SMI results, almost all of the 2012 method-based AMI estimates were smaller than their corresponding 2008 method-based estimates, except for persons aged 50 or older. As with SMI, the 18 to 25 age group had a very large bias measure using the 2008 method. Moreover, most of the 2012 method-based estimates were closer to the direct estimates from the MHSS clinical sample than the 2008 method-based estimates, particularly the AMI estimates for the younger age groups. One subpopulation, the South region, had a significantly biased AMI estimate when applying the 2012 model. However, both the direct estimate for the South and the cut point estimate based on the 2012 model were below the national average.

Table 7.6 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008-2012 Adult NSDUH Main Study and Clinical Sample

Domain	Direct Estimate from Clinical Sample	Estimate Using 2008 Model	Estimate Using 2012 Model
Overall	18.0	20.0	18.1
Age			
18 to 25	20.9	30.5[a]	18.5
26 to 34	19.7	24.1	22.0
35 to 49	20.2	20.5	20.4
50 or Older	14.9	14.4	14.9
Gender			
Male	14.4	16.1	14.4
Female	21.3	23.5	21.5
Race			
White, Not Hispanic	18.2	20.5	19.0
Black, Not Hispanic	15.8	19.4	16.8
Other, Not Hispanic	16.3	19.3	16.8
Hispanic	19.4	17.9	15.3
Region			
Northeast	19.4	19.6	18.0
Midwest	16.4	20.3	18.3
South	16.9	19.7[a]	17.7[a]
West	20.0	20.4	18.4
County Type			
Large Metro	19.3	19.2	17.4
Small Metro	16.9	20.7	18.9
Nonmetro	15.8	20.9[a]	18.6
Received Mental Health Treatment[1]			
Yes	53.2	55.4	53.9
No	12.4	14.3	12.3
Employment Level			
Employed Full Time	15.0	16.6	15.4
Employed Part Time	19.7	23.8	19.3
Unemployed	20.7	28.6	23.8
Other[2]	21.8	22.2[a]	21.0
Education			
Less than High School	25.9	22.5	19.9
High School Graduate	17.2	19.8	17.7
Some College	16.4	21.9[a]	19.6
College Graduate	16.7	17.0	16.1

See notes at end of table. (continued)

Table 7.6 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008-2012 Adult NSDUH Main Study and Clinical Sample (continued)

Domain	Direct Estimate from Clinical Sample	Estimate Using 2008 Model	Estimate Using 2012 Model
Poverty Threshold[3]			
< 100% Threshold	25.5	29.9	25.6
100% - 199% Threshold	24.3	23.6	20.9
≥ 200% Threshold	14.9	16.8	15.7
Health Insurance			
Yes	16.9	19.0	17.4
No	23.4	24.8	21.4

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health.

NOTE: The weights were *not* scaled for the cut point estimators when combining data across years. For the direct estimator, the scaling factors were .06 for the 2008A and 2008B clinical samples, .04 for the 2009 sample, .14 for the 2010 sample, and .35 for the 2011 and 2012 samples.

[a] The difference between the model-based and direct estimates computed from the subclinical sample is significantly different from zero at the .05 level.

[1] Received Mental Health Treatment is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.

[2] Other Employment includes students, persons keeping house or caring for children full time, retired or disabled persons, or other persons not in the labor force.

[3] Poverty Threshold incorporates information on family income, size, and composition and is calculated as a percentage of the U.S. Census Bureau's poverty thresholds. Respondents aged 18 to 22 who were living in a college dormitory were excluded.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; NSDUH main study and clinical sample, 2008-2012.

7.5 Caveats on Model-Based Mental Illness Prevalence Estimates

Various caveats on the methods from producing model-based estimates of SMI should be noted. NSDUH covers only residents of households (i.e., persons living in houses/townhouses, apartments, condominiums; civilians living in housing on military bases, etc.) and persons in noninstitutional group quarters (e.g., shelters, rooming/boarding houses, college dormitories, migratory workers' camps, halfway houses). It does not cover persons who, for the entire year, had no fixed address (e.g., homeless and/or transient persons not in shelters), were on active military duty, or who resided in institutional group quarters (e.g., correctional facilities, nursing homes, mental institutions, and long-term hospitals). Many persons in these excluded categories have mental illness (especially homeless persons and those living in institutional group quarters), but they are not accounted for in the NSDUH SMI and AMI prevalence estimates.

Although effort went into adjusting for the potential biases in those adults responding to the clinical interview, there is no guarantee that the adjustments for nonresponse to the clinical and the undercoverage of Hispanics who chose to respond to the main survey in Spanish were completely successful. There is likewise no guarantee that the nonresponse and coverage adjustments in NSDUH were completely successful in removing all of the biases.

In addition, the mental illness estimates were based on a weighted logistic model for SMI. Although statistical tests did not uncover significant biases in the SMI prevalence estimates within most of the key subpopulations of interest, that does not guarantee that other subpopulation-level biases do not exist. Although many of the subpopulation-level estimates commonly produced were investigated, there are many more possible subpopulations for which SMI estimates may be computed given the vast number of variables available in the NSDUH data. Also, it should be noted that SMI prevalence estimates for a subpopulation closely related to a variable used in the SMI model is likely to be biased. Such variables include suicidal thoughts, the experience of MDE, and the various components of the K6 or WHODAS scales. An ongoing research study is evaluating several alternate models that do not use suicidal thoughts and the experience of MDE as predictors and may produce unbiased estimates for the corresponding subpopulations.

Another caveat concerns the SEs of the SMI estimates. By treating model-predicted SMI indicators as true values of SMI when estimating the SE of a cut point estimate, SEs are underestimated. This underestimation is due to not accounting for the error in model fitting. Nevertheless, SEs calculated in this way can be useful when estimating the difference in SMI prevalence between subpopulations or differences in estimates over time because the same model fit is used for both subpopulations; therefore, the error due to model fitting when estimating differences is effectively cancelled out.

Finally, the MHSS clinical data collection ended in 2012, and the 2012 SMI model will be applied to NSDUH data going forward to produce mental illness estimates under the assumption that the relationship between SMI and the predictor variables is stable over time. Without continuing the clinical interview, the validity of this assumption cannot be assessed, and one cannot test whether covariates other than those specified in the 2012 model may be needed in the future to reduce misclassification or subpopulation-level biases.

References

Aldworth, J., Barnett-Walker, K., Chromy, J., Karg, R., Kott, P., & Morton, K. (2010, December). *Measuring serious mental illness with the NSDUH: Results of 2009 12-month analysis* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022, Mental Health Surveillance Survey Deliverable No. 7, RTI/0209009.523.006.008). Research Triangle Park, NC: RTI International.

Aldworth, J., Barnett-Walker, K., Chromy, J., Karg, R., Morton, K., Novak, S., & Spagnola, K. (2009, June). *Measuring serious mental illness with the NSDUH: Results of 2008 12-month analysis* (prepared for the Substance Abuse and Mental Health Services Administration under Contract No. 283-2004-00022, Mental Health Surveillance Survey Deliverable No. 5, RTI/0209009.423.006.008). Research Triangle Park, NC: RTI International.

Aldworth, J., Kott, P., Yu, F., Mosquin, P., & Barnett-Walker, K. (2012). Analysis of effects of 2008 NSDUH questionnaire changes: Methods to adjust adult MDE and SPD estimates and to estimate SMI in the 2005-2009 surveys. In *2010 National Survey on Drug Use and Health: Methodological resource book* (Section 16b, prepared for the Substance Abuse and Mental Health Services Administration under Contract No. HHSS283200800004C, Deliverable No. 39, RTI/0211838.108.005). Research Triangle Park, NC: RTI International.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (DSM-IV) (4th ed.). Washington, DC: Author.

American Psychiatric Publishing, Inc. (2000). *Diagnostic and statistical manual of mental disorders, 4th ed., text revision* (DSM-IV-TR; doi: 10.1176/appi.books.9780890423349). Retrieved from http://www.psychiatryonline.com/resourceTOC.aspx?resourceID=1

Center for Behavioral Health Statistics and Quality. (2010). *Results from the 2009 National Survey on Drug Use and Health: Mental health findings* (HHS Publication No. SMA 10-4609, NSDUH Series H-39). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2011). *Results from the 2010 National Survey on Drug Use and Health: Summary of national findings* (HHS Publication No. SMA 11-4658, NSDUH Series H-41). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2012a). *Results from the 2010 National Survey on Drug Use and Health: Mental health findings* (HHS Publication No. SMA 11-4667, NSDUH Series H-42). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2012b). *Results from the 2011 National Survey on Drug Use and Health: Summary of national findings* (HHS Publication No. SMA 12-4713, NSDUH Series H-44). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2012c). *Results from the 2011 National Survey on Drug Use and Health: Mental health findings* (HHS Publication No. SMA 12-4725, NSDUH Series H-45). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. (2013). *National Survey on Drug Use and Health: 2012 public use file and codebook* (doi:10.3886/ICPSR34933.v1). Retrieved from http://www.icpsr.umich.edu/icpsrweb/SAMHDA/studies/34933

Center for Behavioral Health Statistics and Quality. (2014a). Mental Health Surveillance Study operations report (2008-2012). In *2012 National Survey on Drug Use and Health: Methodological resource book* (Section 16a). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. (2014b). *National Survey on Drug Use and Health: 2008-2012 adult clinical interview data file and codebook.* Retrieved from http://www.icpsr.umich.edu/icpsrweb/content/SAMHDA/dataportal.html

Center for Behavioral Health Statistics and Quality. (in press a). *Estimating mental illness among adults in the United States: Revisions to 2008 estimation procedures*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (in press b). *Evaluation of imputation methods for the NSDUH*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Chen, P., Cribb, D., Dai, L., Gordek, H., Laufenberg, J., Sathe, N., & Westlake, M. (2014). Person-level sampling weight calibration. In *2012 National Survey on Drug Use and Health: Methodological resource book* (Section 12, prepared for the Substance Abuse and Mental Health Services Administration under Contract No. HHSS283201000003C, Phase I, Deliverable No. 41, RTI/0212800.001.107.004). Research Triangle Park, NC: RTI International.

Clark, W. B., & Hill, M. H. (Eds.). (1991). *Alcohol in America: Drinking practices and problems*. Albany, NY: State University of New York Press.

Colpe, L. J., Barker, P. R., Karg, R. S., Batts, K. R., Morton, K. B., Gfroerer, J. C., Stolzenberg, S. J., Cunningham, D. B., First, M. B., & Aldworth, J. (2010). The National Survey on Drug Use and Health Mental Health Surveillance Study: Calibration study design and field procedures. *International Journal of Methods in Psychiatric Research, 19*(Suppl. 1), 36-48. doi: 10.1002/mpr.311

Crippa, J. A., de Lima Osório, F., Del-Ben, C. M., Filho, A. S., da Silva Freitas, M. C., & Loureiro, S. R. (2008). Comparability between telephone and face-to-face structured clinical interview for DSM-IV in assessing social anxiety disorder. *Perspectives in Psychiatric Care, 44*, 241-247.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33*, 766-771.

Fennig, S., Craig, T., Lavelle, J., Kovasznay, B., & Bromet, E. J. (1994). Best-estimate versus structured interview-based diagnosis in first-admission psychosis. *Comprehensive Psychiatry, 35,* 341-348.

First, M. B., Spitzer R. L., Gibbon M., & Williams J. B. W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition. (SCID-I/NP).* New York, NY: New York State Psychiatric Institute, Biometrics Research Department.

Furukawa, T. A., Kessler, R. C., Slade, T., & Andrews, G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychological Medicine, 33*, 357-362.

Greenfield, T. K., & Kaskutas, L. A. (1998). Five years' exposure to alcohol warning label messages and their impacts: Evidence from diffusion analysis. *Applied Behavioral Science Review, 6*, 39-68.

Hajebi, A., Motevalian, A., Amin-Esmaeili, M., Hefazi, M., Radgoodarzi, R., Rahimi-Movaghar, A., & Sharifi, V. (2012). Telephone versus face-to-face administration of the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, for diagnosis of psychotic disorders. *Comprehensive Psychiatry, 53*, 579-583.

Jackson, J. S., Neighbors, H. W., Nesse, R. M., Trierweller, S. J., & Torres, M. (2004). Methodological innovations in the National Survey of American Life. *International Journal of Methods in Psychiatric Research, 13*, 289-298.

Jordan, B. K., Karg, R. S., Batts, K. R., Epstein, J. F., & Wiesen, C. A. (2008). A clinical validation of the National Survey on Drug Use and Health Assessment of Substance Use Disorders. *Addictive Behaviors, 33*, 782-798.

Kendler, K.S., Neale, M.C., Kessler, R.C., Heath, A.C., & Eaves, L.J. (1992). A population based twin study of major depression in women: The impact of varying definitions of illness. *Archives of General Psychiatry, 49*, 257-266.

Kessler, R., Abelson, J., Demler, O., Escobar, J. I., Gibbon, M., Guyer, M. E., Howes, M. J., Jin, R., Vega, W. A., Walters, E. E., Wang, P., Zaslavsky, A., & Zheng, H. (2004). Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMH-CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 122-139.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. L., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*, 184-189. doi: yoa20567 [pii]

Kott, P. (2011). A nearly pseudo-optimal method for keeping calibration weights from falling below unity in the absence of nonresponse or frame errors. *Pakistan Journal of Statistics, 27*, 391-396.

Kranzler, H. R., Kadden, R. M., Babor, T. F., Tennen, H., & Rounsaville, B. J. (1996). Validity of the SCID in substance abuse patients. *Addiction, 91,* 859-868.

Kranzler, H. R., Kadden, R. M., Burleson, J. A., Babor, T. F., Apter, A., & Rounsaville, B. J. (1995). Validity of psychiatric diagnoses in patients with substance use disorders: Is the interview more important than the interviewer? *Comprehensive Psychiatry, 36*, 278-288.

Lee, S., Tsang, A., Lau, L., Mak, A., Ng, K. L., & Chan, D. M. (2008). Concordance between telephone survey classification and face-to-face structured clinical interview in the diagnosis of generalized anxiety disorder in Hong Kong. *Journal of Anxiety Disorders, 22*, 1403-1411.

Leon, A. C., Olfson, M., Portera, L., Farber, L., & Sheehan, D. V. (1997). Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *International Journal of Psychiatry in Medicine, 27*(2), 93-105.

Liao, D., Aldworth, J., Yu, F., Morton, K., Chen, P., Shook-Sa, B., Kott, P., Davis, T., & Karg, R. (2012). *2011 Mental Health Surveillance Study: Design and estimation report* (prepared for the Substance Abuse and Mental Health Services Administration, Contract No. HHSS283200800004C, RTI 0211838.212.008). Research Triangle Park, NC: RTI International.

Lohr, S. L. (1999). *Sampling: Design and analysis.* Belmont, CA: Duxbury Press.

Maddahian, E., Newcomb, M. D., & Bentler, P. M. (1988). Adolescent drug use and intention to use drugs: Concurrent and longitudinal analyses of four ethnic groups. *Addictive Behaviors, 13*, 191-195.

Morton, K. B., Martin, P. C., Shook-Sa, B. E., Chromy, J. R., & Hirsch, E. L. (2012, June). Sample design report. In *2011 National Survey on Drug Use and Health: Methodological resource book* (Section 2, prepared for the Substance Abuse and Mental Health Services Administration, under Contract No. HHSS283200800004C, Phase II, Deliverable No. 8, RTI/0211838.203.004). Research Triangle Park, NC: RTI International.

Morton, K. B., Martin, P. C., Shook-Sa, B. E., Chromy, J. R., & Hirsch, E. L. (2013, January). Sample design report. In *2012 National Survey on Drug Use and Health: Methodological resource book* (Section 2, prepared for the Substance Abuse and Mental Health Services Administration under Contract No. HHSS283201000003C, Phase I, Deliverable No. 8, RTI/0212800.001.103). Research Triangle Park, NC: RTI International.

Novak, S. P., Colpe, L. J., Barker, P. R., & Gfroerer, J. C. (2010). Development of a brief mental health impairment scale using a nationally representative sample in the USA. *International Journal of Methods in Psychiatric Research*, *19*(Suppl. 1), 49-60. doi:10.1002/mpr.313

Office of Applied Studies. (2009). *Results from the 2008 National Survey on Drug Use and Health: National findings* (HHS Publication No. SMA 09-4434, NSDUH Series H-36). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Office of Applied Studies. (2010a). *Results from the 2009 National Survey on Drug Use and Health: Volume I. Summary of national findings* (HHS Publication No. SMA 10-4586Findings, NSDUH Series H-38A). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Office of Applied Studies. (2010b). *Results from the 2009 National Survey on Drug Use and Health: Volume II. Technical appendices and selected prevalence tables* (HHS Publication No. SMA 10-4586Appendices, NSDUH Series H-38B). Rockville, MD: Substance Abuse and Mental Health Services Administration.

Ramirez Basco, M., Bostic, J. Q., Davies, D., Rush, A. J., Witte, B., Hendrickse, W., & Barnett, V. (2000). Methods to improve diagnostic accuracy in a community mental health setting. *American Journal of Psychiatry, 157*, 1599-1605.

Rehm, J., Üstün, T. B., Saxena, S., Nelson, C. B., Chatterji, S., Ivis, F., & Adlaf, E. (1999). On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. *International Journal of Methods in Psychiatric Research, 8*(2), 110-123. doi:10.1002/mpr.61

Rohde, P., Lewinsohn, P.M., & Seeley, J.R. (1997). Comparability of telephone and face-to-face interviews in assessing axis I and axis II disorders. *American Journal of Psychiatry, 154*, 1593-1598.

Ross, M., Xun, W. Q. E., & Wilson, A. (2002). Language and the bicultural self. *Personality and Social Psychology Bulletin, 28*, 1040-1050.

RTI International. (2012). *SUDAAN®, Release 11.0* [computer software]. Research Triangle Park, NC: Author.

Segal, D. L., Kabacoff, R. I., Hersen, M., Van Hasselt, V. B., & Ryan, C. F. (1995). Update on the reliability of diagnosis in older psychiatric outpatients using the Structured Clinical Interview for DSM-III-R. *Journal of Clinical Geropsychology, 1,* 313-321.

Shear, M. K., Greeno, C., Kang, J., Ludewig, D., Frank, E., Swartz, H. A., & Hanekamp, M. (2000). Diagnosis of nonpsychotic patients in community clinics. *American Journal of Psychiatry, 157*, 581-587.

Silva de Crane, R., & Spielberger, C. D. (1981). Attitudes of Hispanic, black and Caucasian university students toward mental illness. *Hispanic Journal of Behavioral Sciences, 3*, 241-255.

Sobin, C., Weissman, M. M., Goldstein, R. B., Adams, P., Wickramaratne, P., Warner, V., & Lish, J. D. (1993). Diagnostic interviewing for family studies: Comparing telephone and face-to-face methods for the diagnosis of lifetime psychiatric disorders. *Psychiatric Genetics, 3*, 227-233.

Steiner, J. L., Tebes, J. K., Sledge, W. H., & Walker, M. L. (1995). A comparison of the structured clinical interview for DSM-III-R and clinical diagnoses. *Journal of Nervous and Mental Disease, 183,* 365-369.

Substance Abuse and Mental Health Services Administration, Center for Mental Health Services. (1993, May 20). Final notice [Final definitions for: (1) Children with a serious emotional disturbance, and (2) adults with a serious mental illness]. *Federal Register, 58*(96), 29422-29425.

Trafimow, D., Silverman, E. S., Mei-Tai Fan, R., & Shui Fun Law, J. (1997). The effect of language and priming on the relative accessibility of the private self and the collective self. *Journal of Cross-Cultural Psychology, 28*, 107-123.

Triandis, H. C., Davis, E. E., Vassiliou, V., & Nassiakou, M. (1965). *Some methodological problems concerning research on negotiations between monolinguals*. Urbana, IL: University of Illinois.

Zanarini, M. C., & Frankenburg, F. R. (2001). Attainment and maintenance of reliability of axis I and II disorders over the course of a longitudinal study. *Comprehensive Psychiatry, 42,* 369-374.

Zanarini, M. C, Skodol, A. E., Bender, D., Dolan, R., Sanislow, C., Schaefer, E., Morey, L. C., Grilo, C. M., Shea, M. T., McGlashan, T. H., & Gunderson, J. G. (2000). The Collaborative Longitudinal Personality Disorders Study: Reliability of axis I and II diagnoses. *Journal of Personality Disorders, 14,* 291-299.

Appendix A: Kessler-6 (K6) Module

The Kessler-6 (K6) screening instrument for nonspecific psychological distress (Furukawa, Kessler, Slade, & Andrews, 2003; Kessler et al., 2003) were included in the National Survey on Drug Use and Health (NSDUH) main interview. The two six-item K6 scales gather information regarding how frequently a respondent experienced symptoms of psychological distress during the past 30 days and during a month in the past 12 months, respectively. Only respondents who indicated that there was a worse month (DSTWORST=1) in the past 12 months when he or she felt more depressed, anxious, or emotionally stressed than in the past 30 days were asked about the worst month in the past year scale other than the past 30-day scale.

The questions comprising the two K6 scales and the screener question for the worst month scale are provided below with their associated edited variable names from the mental health module, as well as the response categories for each question:

DSTNRV30 During the past 30 days, how often did you feel nervous?

1 All of the time
2 Most of the time
3 Some of the time
4 A little of the time
5 None of the time
DK/REF

Response categories are the same for the remaining past month K6 questions:

DSTHOP30 During the past 30 days, how often did you feel hopeless?

DSTRST30 During the past 30 days, how often did you feel restless or fidgety?

DSTCHR30 During the past 30 days, how often did you feel so sad or depressed that nothing could cheer you up?

DSTEFF30 During the past 30 days, how often did you feel that everything was an effort?

DSTNGD30 During the past 30 days, how often did you feel down on yourself, no good, or worthless?

DSTWORST The last questions asked about how you have been feeling during the past 30 days. Now think about **the past 12 months.** Was there a month in the past 12 months when you felt more depressed, anxious, or emotionally stressed than you felt during the past 30 days?

1 Yes
2 No

DSTNRV12 Think of one month in the past 12 months when you were the most depressed, anxious, or emotionally stressed.

During that month, how often did you feel nervous?

1 All of the time
2 Most of the time
3 Some of the time
4 A little of the time
5 None of the time
DK/REF

Response categories are the same for the remaining worst month K6 questions:

DSTHOP12 During that same month when you were at your worst emotionally . . . how often did you feel hopeless?

DSTRST12 During that same month when you were at your worst emotionally . . . how often did you feel restless or fidgety?

DSTCHR12 During that same month when you were at your worst emotionally . . . how often did you feel so sad or depressed that nothing could cheer you up?

DSTEFF12 During that same month when you were at your worst emotionally . . . how often did you feel that everything was an effort?

DSTNGD12 During that same month when you were at your worst emotionally . . . how often did you feel down on yourself, no good, or worthless?

Each K6 scale item shown above was transformed so that "All of the time" was coded 4, "Most of the time" was coded 3, "Some of the time" was coded 2, "A little of the time" was coded 1, and "None of the time" was coded 0, along with responses matching "Don't know," refusals, bad data, blanks, and legitimate skips.

A past month K6 total score (K6SCMON) was calculated by summing these transformed values across the six past 30-day variables (DSTNRV30, DSTHOP30, DSTRST30, DSTCHR30, DSTEFF30, and DSTNGD30) to arrive at a value ranging between 0 and 24. Likewise, a worst month in the past year K6 total score (K6SCYR) was calculated by summing the transformed values across the six variables for worst month in the past year (DSTNRV12, DSTHOP12, DSTRST12, DSTCHR12, DSTEFF12, and DSTNGD12) to arrive at a value ranging between 0 and 24. The worst month in the past year K6 total score (K6SCYR) has values only for adult respondents who indicated that there was a month in the past year that was worse than the past 30 days (DSTWORST=1). The worst K6 total score in the past year (K6SCMAX) was then created that takes on the maximum value of the past month K6 total score (K6SCMON) and the worst month in the past year K6 total score (K6SCYR) in order to represent the worst K6 total score during the past year, regardless of whether this contradicts the response to DSTWORST.

An alternative past year K6 total score (WSPDSC2) was formulated as follows: past year K6 total scores (K6SCMAX) of less than 8 were recoded as 0, and past year K6 total scores (K6SCMAX) from 8 to 24 were recoded as 1 to 17. The reason for the recoded version was that the serious mental illness (SMI) prevalence was typically extremely low for respondents with

past year K6 total scores of less than 8, and the SMI prevalence rates were higher, in general, only for total scores of 8 or greater. Hence, this variable was used in both the 2008 and 2012 models to create the mental illness variables.

Appendix B: World Health Organization Disability Assessment Schedule (WHODAS) Module

The World Health Organization Disability Assessment Schedule (WHODAS) is a scale used to measure functional impairment that consists of a series of items that are used for assessing disturbances in social adjustment and behavior (i.e., functional impairment). A reduced set of 13 WHODAS items (Novak, Colpe, Barker, & Gfroerer, 2010; Rehm et al., 1999) were administered to a random half of the adult sample in the 2008 National Survey on Drug Use and Health (NSDUH) and to the total adult sample in subsequent NSDUHs (starting in 2009). Responses to this impairment scale were used to create eight variables that were transformed and summed to define the WHODAS total score used in the development of both the 2008 and 2012 serious mental illness (SMI) prediction models. The questions comprising the abbreviated WHODAS are provided below with their associated edited variable names from the mental health module, as well as the response categories for each question:

The next questions are about how much your emotions, nerves, or mental health caused you to have **difficulties in daily activities**.

In answering, think of the **one month** in the past 12 months when your emotions, nerves, or mental health interfered **most** with your daily activities.

IMPREMEM During that one month when your emotions, nerves or mental health interfered **most** with your daily activities . . .

how much difficulty did you have **remembering to do things you needed to do**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
DK/REF

IMPCONCN how much difficulty did you have **concentrating on doing something important when other things were going on around you**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
DK/REF

IMPGOUT how much difficulty did you have **going out of the house and getting around on your own**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
5 You didn't leave the house on your own
DK/REF

IMPGOUTM [IF IMPGOUT = 5] Did problems with your emotions, nerves, or mental health keep you from leaving the house on your own?

1 Yes
2 No
DK/REF

IMPPEOP how much difficulty did you have **dealing with people you did not know well**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
5 You didn't deal with people you did not know well
DK/REF

IMPPEOPM [IF IMPPEOP = 5] Did problems with your emotions, nerves, or mental health keep you from dealing with people you did not know well?

1 Yes
2 No
DK/REF

IMPSOC how much difficulty did you have **participating in social activities, like visiting friends or going to parties**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
5 You didn't participate in social activities
DK/REF

IMPSOCM [IF IMPSOC=5] Did problems with your emotions, nerves, or mental health keep you from participating in social activities?

1 Yes
2 No
DK/REF

IMPHHLD how much difficulty did you have **taking care of household responsibilities**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
5 You didn't take care of household responsibilities
DK/REF

IMPHHLDM [IF IMPHHLD=5] Did problems with your emotions, nerves, or mental health keep you from taking care of household responsibilities?

1 Yes
2 No
DK/REF

IMPRESP how much difficulty did you have **taking care of your daily responsibilities at work or school**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
5 You didn't work or go to school
DK/REF

IMPRESPM [IF IMPRESP=5] Did problems with your emotions, nerves, or mental health keep you from working or going to school?

1 Yes
2 No
DK/REF

IMPWORK [IF **IMPRESP** NE 5] During that one month when your emotions, nerves or mental health interfered **most** with your daily activities . . .

how much difficulty did you have **getting your daily work done as quickly as needed**?

1 No difficulty
2 Mild difficulty
3 Moderate difficulty
4 Severe difficulty
DK/REF

An original WHODAS total score (WHODASC2) was created to indicate the level of difficulty in performing daily activities due to problems with emotions, nerves, or mental health. Each of the eight variables created from the WHODAS items shown above was transformed into values of 0 to 3 so that a response of "severe difficulty" was coded 3, "moderate difficulty" was coded 2, "mild difficulty" was coded 1, and "no difficulty" was coded 0, with "don't know" and "refuse" also coded 0. Some items had a fifth category to deal with "not applicable" responses. For example, the question about difficulties regarding taking care of daily responsibilities at work or school (impresp) had a fifth category, "you didn't go to work or school." If this category was selected, then another question was asked as to whether respondents' emotions, nerves, or mental health caused them to be unable to go to work or school (imprespm). A "yes" response to the follow-up question (imprespm=1) was coded 3, and a "no" response (imprespm=2) was coded 0. One exception to this coding was the last WHODAS recode on how much difficulty the respondents had in getting their daily work done as quickly as needed (impwork). This item was asked of the respondents only if in the previous question they responded that they went to work or school (impresp=1 to 4). In the case that they responded that they did not go to work or school (impresp=5), their response to the follow-up question referred to above (imprespm) determined the final item score for impwork; otherwise, impwork was recoded similar to the other items.

The transformed scale values were summed across the eight variables created from the WHODAS items (remembering, concentrating, going out of the house on your own, dealing with people you don't know well, participating in social activities, taking care of household responsibilities, taking care of daily work/school responsibilities, and getting your daily work done as quickly as needed) to arrive the WHODAS total score (WHODASC2) at a value ranging between 0 and 24.

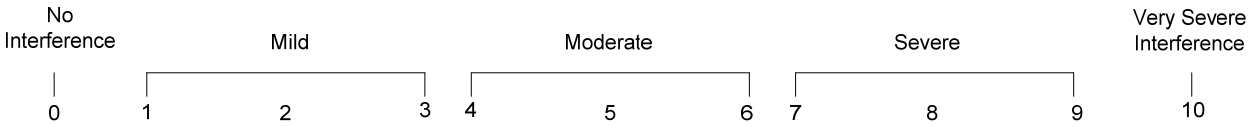
An alternative WHODAS total score (WHODASC3) was created to indicate the number of daily activities in which a respondent had moderate or severe difficulty performing or did not perform due to problems with emotions, nerves, or mental health. Each of the eight variables created from WHODAS items shown above was transformed into values of 0 or 1 so that responses indicating "moderate difficulty" or "severe difficulty" were recoded 1 and responses indicating "mild difficulty" or "no difficulty" were recoded 0. If a fifth category of "not applicable" was available and selected (see above for an example of a fifth question), then another question was asked as to whether respondents' emotions, nerves, or mental health caused them to be unable to go to work or school. A "yes" response was recoded 1, and a "no" response was recoded 0. The transformed scale values were summed across the eight WHODAS activities to arrive at a value ranging between 0 and 8. The recoded version of the WHODAS total score was driven by the idea that a dichotomous measure dividing respondents who experienced moderate or severe difficulties from the remaining respondents might fit better than a linear continuous measure. This alternative WHODAS total score was used in both the 2008 and 2012 models for estimating the mental illness variables.

Appendix C: Sheehan Disability Scale (SDS) Module

The Sheehan Disability Scale (SDS) is a scale used to measure functional impairment that consists of four questions that ask respondents how much their emotions, nerves, or mental health interfered with their daily activities over the past year (Leon, Olfson, Portera, Farber, & Sheehan, 1997). It was administered to a random half of the adult sample in the 2008 National Survey on Drug Use and Health (NSDUH). The questions comprising the SDS are provided below with their associated edited variable names from the mental health module, as well as the response categories for each question:

The next questions are about how much your emotions, nerves, or mental health **interfered with your daily activities**. In answering, think of the **one month** in the past 12 months when your emotions, nerves, or mental health interfered **most** with your daily activities. Using the 0 to 10 scale shown below, where 0 means **no** interference and 10 means very **severe** interference, select the number that describes how much **your emotions, nerves or mental health** interfered with your **ability to do** each of the following activities during that period. You can use any number between 0 and 10 to answer.

MHSMGT **During that month when you were at your worst emotionally,** how much did your emotions interfere with your **ability to do home management tasks**, like cleaning, shopping, and working around the house, apartment, or yard?



Response categories are the same for the remaining SDS questions:

MHSWORK During that **month** in the past 12 months when you were at your worst emotionally how much did this interfere with your **ability to work**?

MHSRELS During that month when you were at your worst emotionally, how much did this interfere with your **ability to form and maintain close relationships** with other people?

MHSSOC How much did your emotions interfere with your **ability to have a social life** during that period of time?

The original SDS total score (SDSSC2) was created to indicate the level of interference with daily activities due to problems with emotions, nerves, or mental health. Each SDS item shown above was transformed so that responses indicating "very severe interference" were coded as 10; responses indicating "severe interference" were coded as 7, 8, and 9; responses indicating "moderate interference" were coded as 4, 5, and 6; responses indicating "mild interference" were coded as 1, 2, and 3; and responses indicating "no interference" were coded as 0. Additionally, sample B respondents who did not answer the questions or those with unknown data were coded as 0. Youths aged 12 to 17 and those adult respondents not in sample B or not assigned to a subsample were assigned the standard SAS missing code (.). The transformed scale values were summed across the four SDS activities (home management, ability to work, ability to form and

maintain close relationships, and ability to have a social life) to arrive at a value ranging between 0 and 40.

The alternative SDS total score (SDSSC3) indicates the number of daily activities in which emotions, nerves, or mental health severely or very severely interfered with performing the activity in the one month in the past year when they were at their worst emotionally. Each SDS item shown above was transformed so that responses indicating "severe interference" or "very severe interference" were recoded as 1. Responses indicating "no interference," "mild interference," or "moderate interference" and sample B respondents who did not answer the questions or those with unknown data were recoded as 0. The transformed scale values were summed across the four SDS activities (see above) to arrive at a value ranging between 0 and 4. The recoded version of the SDS total score also was driven by the idea that a dichotomous measure dividing respondents who experienced severe or very severe interference from the remaining respondents might fit better than a linear continuous measure.

Appendix D: Descriptive Characteristics for Undercoverage of Hispanics, Agreement to Participate, and Completion of the Clinical Follow-Up

Table D1a Descriptive Characteristics, by Language of Interview: Hispanic Respondents Aged 18 or Older, 2008-2011 NSDUH

Variable	Hispanics Responding in English Estimate (%)	Hispanics Responding in Spanish Estimate (%)
GENDER		
Male	51.8[a]	49.7
Female	48.2[a]	50.3
AGE		
18 to 25	25.2[c]	10.5
26 to 34	23.7[b]	21.6
35 to 49	28.6[c]	36.0
50 or Older	22.5[c]	31.9
AGE		
18 to 30	39.3[c]	22.1
31 or older	60.7[c]	77.9
COUNTY TYPE		
Large Metro	64.6[c]	73.0
Small Metro	28.0[c]	22.4
Nonmetro	7.4[c]	4.5
EDUCATION		
Less than High School	21.3[c]	62.5
High School Graduate	32.8[c]	22.2
Some College	28.2[c]	9.4
College Graduate	17.7[c]	5.9
MARITAL STATUS		
Married	45.5[c]	57.8
Widowed	2.5[c]	5.2
Divorced or Separated	14.2[b]	12.4
Never Married	37.8[c]	24.6
INCOME		
Less than $20,000	20.0[c]	37.7
$20,000 to $49,999	39.5[c]	49.0
$50,000 to $74,999	16.3[c]	8.1
$75,000 or More	24.3[c]	5.3
REGION		
Northeast	15.7[c]	12.5
Midwest	9.3[c]	7.5
South	36.8	34.5
West	38.2[c]	45.5
NOT CURRENTLY WORKING BUT WORKED IN THE PAST 12 MONTHS[1]		
Yes	9.1[c]	7.4
Other[2]	90.9[c]	92.6
RETIREMENT STATUS IN PAST WEEK		
No Job: Retired	6.7	8.1
Other[3]	93.3	91.9
WORK SITUATION IN PAST WEEK		
No Job: Looking for Work/Layoff, Not Looking for Work/Disabled for Work	12.5	11.3
Other[4]	87.5	88.7
HEALTH INSURANCE[5]		
Covered by Any Health Insurance	71.1[c]	45.2
Not Covered by Any Health Insurance	28.9[c]	54.8
HAD DEPRESSION IN LIFETIME[6]		
Yes	8.9[c]	5.3
No	91.1[c]	94.7

(continued)

Table D1a Descriptive Characteristics, by Language of Interview: Hispanic Respondents Aged 18 or Older, 2008-2011 NSDUH (continued)

Variable	Hispanics Responding in English Estimate (%)	Hispanics Responding in Spanish Estimate (%)
HAD ULCER(S) IN LIFETIME[6]		
Yes	2.0	2.3
No	98.0	97.7
YEARS LIVED IN THE U.S.		
Less than 5 years	1.5[c]	12.5
At least 5 years but less than 10 years	3.3[c]	18.8
10 years or more	25.8[c]	64.5
Born in U.S.	69.4[c]	4.1
INTERVIEW PRIVACY		
Completely private	82.1	80.6
Minor distractions	13.1[b]	14.9
Person(s) in the room or listening about 1/3 time	2.1	2.0
Serious interruptions of privacy more than 1/2 time	0.4	0.4
Constant presence of other person(s)	2.4	2.1
INTERVIEW LENGTH: Adult Depression		
Less than 2 minutes[7]	83.5[c]	91.3
Greater or equal to 2 minutes	16.5[c]	8.7
INTERVIEW LENGTH: Entire Interview		
Less than 60 minutes[7]	39.9[c]	13.8
Greater or equal to 60 minutes	60.1[c]	86.2
NO DIFFICULTY UNDERSTANDING THE MAIN NSDUH INTERVIEW		
No difficulty	88.3[c]	78.7
Other[8]	11.7[c]	21.3

NSDUH = National Survey on Drug Use and Health.

* Low precision.

NOTE: For 2008, the 10 cases that were in neither sample (A or B) have been excluded.

[a] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.10$.

[b] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.05$.

[c] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.01$.

[1] This question only includes respondents who said DK/REF to having worked in the last week or NO/DK/REF to DID YOU HAVE JOB BUSINESS in the past week.

[2] Other includes legitimate skip, did not work in the past 12 months, or missing.

[3] Other includes other no job categories, has job, or missing.

[4] Other includes other no job categories, has job, or missing.

[5] Respondents with unknown health insurance status were excluded.

[6] Respondents with unknown health data were excluded.

[7] Missing or negative data are included in this category.

[8] Other includes little difficulty, some difficulty, a lot of difficulty, or missing.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2011.

Table D1b Mental Health Measures, by Language of Interview: Hispanic Respondents Aged 18 or Older, 2008-2011 NSDUH

Variable	Hispanics Responding in English Estimate (%)	Hispanics Responding in Spanish Estimate (%)
K6 Total Score		
K6 Total Score: Maximum of Past Year and Past 30 Days (Range: 0-24)[1]	5.1[c]	3.3
Score 0 to 3	51.2[c]	68.3
Score 4 to 5	13.8[c]	9.5
Score 6 to 7	9.7[c]	6.5
Score 8 to 9	6.5[c]	4.7
Score 10 to 11	4.8[c]	3.1
Score 12 to 15	6.7[c]	3.3
Score 16 or higher	7.4[c]	4.6
WHODAS Total Score[2]		
Total Score (Range 0-24)[1]	3.4[c]	1.9
Alternative Total Score (Range 0-8)[1]	0.8[c]	0.4
MDE[3]		
Lifetime MDE	11.6[c]	5.4
Past Year MDE	6.4[c]	3.3
MDE TREATMENT[4]		
Saw or Talk to MD/Professional for MDE in Past Year	53.6	49.5
Used Rx Medication for MDE in Past Year	35.3	*
MENTAL HEALTH TREATMENT		
Received Any Mental Health Treatment in Past Year[5]	9.2[c]	3.9
Stay over in Hospital for Mental Health Treatment in Past 12 Months[5]	1.1[c]	0.6
Received Outpatient Mental Health Treatment in Past 12 Months[5]	5.2[c]	1.7
Needed Mental Health Treatment but Didn't Get It in Past 12 Months	5.0[c]	1.7
SUICIDAL THOUGHTS[6]		
Seriously Think about Killing Self in Past 12 Months	3.6[c]	1.2
Make Plans to Kill Yourself in Past 12 Months	1.0[c]	0.3
Try to Kill Yourself in Past 12 Months	0.6[c]	0.2
SUICIDAL THOUGHTS TREATMENT[6]		
Received Medical Attention Because Tried to Kill Self in Past 12 Months	0.4[c]	0.0
Stay Overnight at Hospital Because Tried to Kill Self in Past 12 Months	0.2[c]	0.0

K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; NSDUH = National Survey on Drug Use and Health; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

* Low precision.

NOTE: For 2008, the 10 cases that were in neither sample (A or B) have been excluded. K6SCMAX is used in this table, which is defined as the higher K6 score between the past month K6 total score and the K6 total score in the worst month of the past year if the worst month was not the past 30 days.

[a] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.10$.
[b] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.05$.
[c] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.01$.
[1] Estimates are weighted means.
[2] For 2008, only sample A data are included.
[3] Respondents with unknown lifetime or past year MDE data were excluded.
[4] Among those with past year MDE.
[5] Mental health treatment/counseling is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.
[6] Respondents with unknown suicide information were excluded.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2011.

Table D1c Substance Use Measures, by Language of Interview: Hispanic Respondents Aged 18 or Older, 2008-2011 NSDUH

Variable	Hispanics Responding in English Estimate (%)	Hispanics Responding in Spanish Estimate (%)
LIFETIME SUBSTANCE USE		
ILLICIT DRUGS[1]	50.1[c]	17.0
Marijuana and Hashish	44.0[c]	8.4
Cocaine	16.8[c]	6.0
Crack	3.5[c]	0.9
Heroin	2.2[c]	0.3
Hallucinogens	15.2[c]	1.1
LSD	8.3[c]	0.2
PCP	2.8[c]	0.1
Ecstasy	7.4[c]	0.4
Inhalants	8.3[c]	1.9
Nonmedical Use of Psychotherapeutics[2,3]	20.7[c]	8.1
Pain Relievers	14.3[c]	6.1
OxyContin®	1.8[c]	0.1
Tranquilizers	7.4[c]	2.1
Stimulants[3]	8.4[c]	1.3
Methamphetamine[3]	5.8[c]	0.8
Sedatives	2.1[c]	0.5
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	32.4[c]	13.7
ALCOHOL	85.8[c]	67.6
TOBACCO[4]	67.7[c]	46.0
CIGARETTES	63.9[c]	44.2
PAST YEAR SUBSTANCE USE		
ILLICIT DRUGS[1]	17.7[c]	5.3
Marijuana and Hashish	13.7[c]	1.4
Cocaine	2.6[c]	1.1
Crack	0.3[b]	0.1
Heroin	0.3[b]	0.1
Hallucinogens	2.1[c]	0.2
LSD	0.3[c]	0.0
PCP	0.1[c]	0.0
Ecstasy	1.6[c]	0.1
Inhalants	0.7[c]	0.2
Nonmedical Use of Psychotherapeutics[2,3]	6.5[c]	3.3
Pain Relievers	5.1[c]	2.5
OxyContin®	0.4[c]	*
Tranquilizers	1.8[c]	0.8
Stimulants[3]	1.1[c]	0.3
Methamphetamine[3]	0.5[c]	0.1
Sedatives	0.3[a]	0.1
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	9.1[c]	4.4
ALCOHOL	72.2[c]	46.4
TOBACCO[4]	34.5[c]	19.9
CIGARETTES	30.5[c]	19.2

(continued)

Table D1c Substance Use Measures, by Language of Interview: Hispanic Respondents Aged 18 or Older, 2008-2011 NSDUH (continued)

Variable	Hispanics Responding in English Estimate (%)	Hispanics Responding in Spanish Estimate (%)
PAST YEAR SUBSTANCE ABUSE OR DEPENDENCE[5]		
ALCOHOL ABUSE OR DEPENDENCE	9.6[c]	5.7
PAIN RELIEVER ABUSE OR DEPENDENCE	0.6	0.4

LSD = lysergic acid diethylamide; NSDUH = National Survey on Drug Use and Health; PCP = phencyclidine.

* Low precision.

NOTE: For 2008, the 10 cases that were in neither sample (A or B) have been excluded.

[a] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.10$.

[b] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.05$.

[c] Difference between this estimate and corresponding estimate from Hispanics Responding in Spanish is statistically significant at $p < 0.01$.

[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. The estimates for Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine incorporated in these summary estimates do not include data from the methamphetamine items added in 2005 and 2006. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.

[2] Nonmedical use of prescription-type psychotherapeutics includes the nonmedical use of pain relievers, tranquilizers, stimulants, or sedatives and does not include over-the-counter drugs.

[3] Estimates of Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine in the designated rows include data from methamphetamine items added in 2005 and 2006 and are not comparable with estimates presented in NSDUH reports prior to the 2007 national findings report. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.

[4] Tobacco Products include cigarettes, smokeless tobacco (i.e., chewing tobacco or snuff), cigars, or pipe tobacco.

[5] Dependence or abuse is based on definitions found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2011.

Table D2a Descriptive Characteristics, by Agreement to Participate in the Clinical Follow-Up, 2008-2011 MHSS

Variable	Persons Agreeing to Participate Estimate (%)	Initial Nonrespondents Estimate (%)
GENDER		
Male	48.3	47.9
Female	51.7	52.1
AGE		
18 to 25	18.4[c]	8.6
26 to 34	16.1[c]	8.6
35 to 49	27.2	31.2
50 or Older	38.3[b]	51.6
AGE		
18 to 30	28.3[c]	13.5
31 or older	71.7[c]	86.5
COUNTY TYPE		
Large Metro	54.8	61.9
Small Metro	30.1[b]	22.3
Nonmetro	15.1	15.8
RACE/ETHNICITY		
White	65.9	68.0
Black	10.2[c]	4.6
Hispanic	17.4	*
Asian and Others	6.5	7.7
EDUCATION		
Less than High School	14.2	*
High School Graduate	33.0	39.0
Some College	26.1[c]	17.9
College Graduate	26.8	27.5
MARITAL STATUS		
Married	51.1[b]	63.8
Widowed	4.1	5.8
Divorced or Separated	14.6	13.1
Never Married	30.3[c]	17.3
INCOME		
Less than $20,000	17.2	11.9
$20,000 to $49,999	33.3	37.0
$50,000 to $74,999	18.6	14.6
$75,000 or More	30.9	36.5
REGION		
Northeast	21.5	22.5
Midwest	21.8	18.0
South	31.7	36.4
West	25.0	23.1
NOT CURRENTLY WORKING BUT WORKED IN THE PAST 12 MONTHS[1]		
Yes	8.1[c]	3.8
Other[2]	91.9[c]	96.2
RETIREMENT STATUS IN PAST WEEK		
No Job: Retired	10.7[b]	22.3
Other[3]	89.3[b]	77.7
WORK SITUATION IN PAST WEEK		
No Job: Looking for Work/Layoff, Not Looking for Work/ Disabled for Work	12.6[c]	3.9
Other[4]	87.4[c]	96.1

(continued)

Table D2a Descriptive Characteristics, by Agreement to Participate in the Clinical Follow-Up, 2008-2011 MHSS (continued)

Variable	Persons Agreeing to Participate Estimate (%)	Initial Nonrespondents Estimate (%)
HEALTH INSURANCE[5]		
Covered by Any Health Insurance	81.4	85.9
Not Covered by Any Health Insurance	18.6	14.1
HAD DEPRESSION IN LIFETIME[6]		
Yes	15.1[c]	5.9
No	84.9[c]	94.1
HAD ULCER(S) IN LIFETIME[6]		
Yes	5.6[b]	0.9
No	94.4[b]	99.1
YEARS LIVED IN THE US		
Less than 5 years	1.7[a]	0.5
At least 5 years but less than 10 years	2.0	4.1
10 years or more	12.7	*
Born in US	83.6	77.6
INTERVIEW PRIVACY		
Completely private	87.8	86.7
Minor distractions	8.9	*
Person(s) in the room or listening about 1/3 time	1.3	0.8
Serious interruptions of privacy more than 1/2 time	0.2	0.3
Constant presence of other person(s)	1.8	1.7
INTERVIEW LENGTH: Adult Depression		
Less than 2 minutes[7]	77.2[c]	92.5
Greater or equal to 2 minutes	22.8[c]	7.5
INTERVIEW LENGTH: Entire Interview		
Less than 60 minutes[7]	27.7[c]	49.9
Greater or equal to 60 minutes	72.3[c]	50.1
NO DIFFICULTY UNDERSTANDING THE MAIN NSDUH INTERVIEW		
No difficulty	90.3	88.0
Other[8]	9.7	12.0

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

* Low precision.

NOTE: Table subset to only those cases where MHSELECT=1.

[a] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.10$.
[b] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.05$.
[c] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.01$.
[1] This question only includes respondents who said DK/REF to having worked in the last week or NO/DK/REF to DID YOU HAVE JOB BUSINESS in the past week.
[2] Other includes legitimate skip, did not work in the past 12 months, or missing.
[3] Other includes other no job categories, has job, or missing.
[4] Other includes other no job categories, has job, or missing.
[5] Respondents with unknown health insurance status were excluded.
[6] Respondents with unknown health data were excluded.
[7] Missing or negative data are included in this category.
[8] Other includes little difficulty, some difficulty, a lot of difficulty, or missing.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table D2b Mental Health Measures, by Agreement to Participate in the Clinical Follow-Up, 2008-2011 MHSS

Variable	Persons Agreeing to Participate Estimate (%)	Initial Nonrespondents Estimate (%)
K6 Total Score		
K6 Total Score: Maximum of Past Year and Past 30 Days (Range: 0-24)[1]	5.2[c]	3.2
Score 0 to 3	51.6[c]	66.9
Score 4 to 5	12.8	12.8
Score 6 to 7	10.6[b]	6.5
Score 8 to 9	6.3	4.5
Score 10 to 11	4.1[b]	2.6
Score 12 to 15	7.3[c]	3.6
Score 16 or higher	7.2[c]	3.1
WHODAS Total Score[2]		
Total Score (Range 0-24)[1]	3.6[c]	2.4
Alternative Total Score (Range 0-8)[1]	0.9[c]	0.6
MDE[3]		
Lifetime MDE	14.7[c]	5.4
Past Year MDE	7.4[c]	3.1
MDE TREATMENT[4]		
Saw or Talk to MD/Professional for MDE in Past Year	64.9	*
Used Rx Medication for MDE in Past Year	58.2	*
MENTAL HEALTH TREATMENT		
Received Any Mental Health Treatment in Past Year[5]	15.2[c]	8.7
Stay over in Hospital for Mental Health Treatment in Past 12 Months[5]	0.7[c]	0.1
Received Outpatient Mental Health Treatment in Past 12 Months[5]	7.7[a]	4.9
Needed Mental Health Treatment but Didn't Get It in Past 12 Months	5.7[c]	1.8
SUICIDAL THOUGHTS[6]		
Seriously Think about Killing Self in Past 12 Months	4.2[c]	1.3
Make Plans to Kill Yourself in Past 12 Months	1.1[c]	0.3
Try to Kill Yourself in Past 12 Months	0.4	0.2
SUICIDAL THOUGHTS TREATMENT[6]		
Received Medical Attention Because Tried to Kill Self in Past 12 Months	0.3[a]	0.1
Stay Overnight at Hospital Because Tried to Kill Self in Past 12 Months	0.2	0.1

K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; MHHS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; WHODAS = 8-item World Health Organization Disability Assessment Schedule.
*Low precision.

NOTE: Table subset to only those cases where MHSELECT=1. K6SCMAX is used in this table, which is defined as the higher K6 score between the past month K6 total score and the K6 total score in the worst month of the past year if the worst month was not the past 30 days.

[a] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.10$.
[b] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.05$.
[c] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.01$.
[1] Estimates are weighted means.
[2] For 2008, only sample A data are included.
[3] Respondents with unknown lifetime or past year MDE data were excluded.
[4] Among those with past year MDE.
[5] Mental health treatment/counseling is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.
[6] Respondents with unknown suicide information were excluded.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table D2c Substance Use Measures, by Agreement to Participate in the Clinical Follow-Up, 2008-2011 MHSS

Variable	Persons Agreeing to Participate Estimate (%)	Initial Nonrespondents Estimate (%)
LIFETIME		
ILLICIT DRUGS[1]	57.1[c]	38.7
Marijuana and Hashish	53.0[c]	35.4
Cocaine	19.6[c]	10.4
Crack	4.1[c]	1.5
Heroin	2.7[c]	0.8
Hallucinogens	20.8[c]	6.6
LSD	14.3[c]	3.8
PCP	3.1[c]	0.8
Ecstasy	7.7[c]	1.6
Inhalants	11.1[c]	4.3
Nonmedical Use of Psychotherapeutics[2,3]	24.9[c]	13.7
Pain Relievers	18.9[c]	6.1
OxyContin®	3.2[b]	1.5
Tranquilizers	12.9	7.8
Stimulants[3]	11.3[c]	3.3
Methamphetamine[3]	7.7[c]	1.6
Sedatives	4.3[c]	1.2
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	35.5[c]	23.0
ALCOHOL	88.5[a]	78.8
TOBACCO[4]	74.3[b]	61.8
CIGARETTES	67.6	59.5
PAST YEAR		
ILLICIT DRUGS[1]	19.2[c]	6.8
Marijuana and Hashish	15.1[c]	4.3
Cocaine	2.7[c]	0.4
Crack	0.5[c]	0.0
Heroin	0.3[c]	0.0
Hallucinogens	1.8[c]	0.6
LSD	0.3[b]	0.0
PCP	0.0	*
Ecstasy	1.3[c]	0.3
Inhalants	0.4[a]	0.1
Nonmedical Use of Psychotherapeutics[2,3]	9.4[c]	3.5
Pain Relievers	5.7[c]	2.5
OxyContin®	0.8	0.6
Tranquilizers	4.6	1.8
Stimulants[3]	3.3	0.3
Methamphetamine[3]	*	0.1
Sedatives	0.5[a]	0.1
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	11.8[c]	4.2
ALCOHOL	71.7	69.2
TOBACCO[4]	36.7[b]	26.5
CIGARETTES	31.1[b]	22.8

(continued)

Table D2c Substance Use Measures, by Agreement to Participate in the Clinical Follow-Up, 2008-2011 MHSS (continued)

Variable	Persons Agreeing to Participate Estimate (%)	Initial Nonrespondents Estimate (%)
PAST YEAR SUBSTANCE ABUSE OR DEPENDENCE[5]		
ALCOHOL ABUSE OR DEPENDENCE	8.3[b]	4.8
PAIN RELIEVER ABUSE OR DEPENDENCE	1.1	0.6

LSD = lysergic acid diethylamide; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; PCP = phencyclidine.

* Low precision.

NOTE: Table subset to only those cases where MHSELECT=1.

[a] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.10$.
[b] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.05$.
[c] Difference between this estimate and corresponding estimate from initial nonrespondents is statistically significant at $p < 0.01$.
[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. The estimates for Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine incorporated in these summary estimates do not include data from the methamphetamine items added in 2005 and 2006. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.
[2] Nonmedical use of prescription-type psychotherapeutics includes the nonmedical use of pain relievers, tranquilizers, stimulants, or sedatives and does not include over-the-counter drugs.
[3] Estimates of Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine in the designated rows include data from methamphetamine items added in 2005 and 2006 and are not comparable with estimates presented in NSDUH reports prior to the 2007 national findings report. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.
[4] Tobacco Products include cigarettes, smokeless tobacco (i.e., chewing tobacco or snuff), cigars, or pipe tobacco.
[5] Dependence or abuse is based on definitions found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table D3a Descriptive Characteristics, by Participation in Clinical Interview: Persons Initially Agreeing to Participate in Clinical Follow-Up, 2008-2011 MHSS

Variable	Respondents to Clinical Follow-Up Estimate (%)	Final Nonrespondents Estimate (%)
GENDER		
Male	46.4[c]	60.6
Female	53.6[c]	39.4
AGE		
18 to 25	14.3[b]	22.6
26 to 34	15.3	19.0
35 to 49	25.6	32.8
50 or Older	44.8[c]	*
AGE		
18 to 30	23.0[c]	36.6
31 or older	77.0[c]	63.4
COUNTY TYPE		
Large Metro	52.1	52.2
Small Metro	29.6	35.7
Nonmetro	18.3[b]	12.0
RACE/ETHNICITY		
White	73.3[c]	55.2
Black	8.4	10.5
Hispanic	12.4[b]	*
Asian and Others	5.8	6.5
EDUCATION		
Less than High School	9.4[c]	*
High School Graduate	35.5	30.9
Some College	25.9	21.8
College Graduate	29.2[c]	16.1
MARITAL STATUS		
Married	56.6[a]	44.9
Widowed	4.3	*
Divorced or Separated	14.7	11.7
Never Married	24.4[c]	39.8
INCOME		
Less than $20,000	13.7[b]	22.9
$20,000 to $49,999	33.3	34.8
$50,000 to $74,999	20.5[c]	12.5
$75,000 or More	32.5	29.8
REGION		
Northeast	24.2[a]	17.3
Midwest	24.1[a]	17.1
South	31.0	35.3
West	20.7[a]	30.4
NOT CURRENTLY WORKING BUT WORKED IN THE PAST 12 MONTHS[1]		
Yes	7.7	5.4
Other[2]	92.3	94.6
RETIREMENT STATUS IN PAST WEEK		
No Job: Retired	14.9[a]	7.9
Other[3]	85.1[a]	92.1
WORK SITUATION IN PAST WEEK		
No Job: Looking for Work/Layoff, Not Looking for Work/Disabled for Work	11.4	12.0
Other[4]	88.6	88.0

(continued)

Table D3a Descriptive Characteristics, by Participation in Clinical Interview: Persons Initially Agreeing to Participate in Clinical Follow-Up, 2008-2011 MHSS (continued)

Variable	Respondents to Clinical Follow-Up Estimate (%)	Final Nonrespondents Estimate (%)
HEALTH INSURANCE[5]		
Covered by Any Health Insurance	86.0[c]	69.6
Not Covered by Any Health Insurance	14.0[c]	30.4
HAD DEPRESSION IN LIFETIME[6]		
Yes	13.0	13.0
No	87.0	87.0
HAD ULCER(S) IN LIFETIME[6]		
Yes	3.6	*
No	96.4	*
YEARS LIVED IN THE US		
Less than 5 years	1.0	*
At least 5 years but less than 10 years	1.5	3.3
10 years or more	10.5	*
Born in US	87.0	*
INTERVIEW PRIVACY		
Completely private	88.8	85.3
Minor distractions	8.1	10.7
Person(s) in the room or listening about 1/3 time	1.4	0.9
Serious interruptions of privacy more than 1/2 time	0.1	0.2
Constant presence of other person(s)	1.6	2.9
INTERVIEW LENGTH: Adult Depression		
Less than 2 minutes[7]	81.7	77.3
Greater or equal to 2 minutes	18.3	22.7
INTERVIEW LENGTH: Entire Interview		
Less than 60 minutes[7]	31.7	36.3
Greater or equal to 60 minutes	68.3	63.7
NO DIFFICULTY UNDERSTANDING THE MAIN NSDUH INTERVIEW		
No difficulty	93.7[a]	*
Other[8]	6.3[a]	*

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

* Low precision.

NOTE: Table subset to only those cases who initially agreed to participate in the clinical follow-up interview.

[a] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.10$.
[b] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.05$.
[c] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.01$.
[1] This question only includes respondents who said DK/REF to having worked in the last week or NO/DK/REF to DID YOU HAVE JOB BUSINESS in the past week.
[2] Other includes legitimate skip, did not work in the past 12 months, or missing.
[3] Other includes other no job categories, has job, or missing.
[4] Other includes other no job categories, has job, or missing.
[5] Respondents with unknown health insurance status were excluded.
[6] Respondents with unknown health data were excluded.
[7] Missing or negative data are included in this category.
[8] Other includes little difficulty, some difficulty, a lot of difficulty, or missing.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table D3b Mental Health Measures, by Participation in Clinical Interview: Persons Initially Agreeing to Participate in Clinical Follow-Up, 2008-2011 MHSS

Variable	Respondents to Clinical Follow-Up Estimate (%)	Final Nonrespondents Estimate (%)
K6 Total Score		
K6 Total Score: Maximum of Past Year and Past 30 Days (Range: 0-24)[1]	4.7	4.9
Score 0 to 3	54.3	56.6
Score 4 to 5	13.6	10.0
Score 6 to 7	11.3[b]	7.4
Score 8 to 9	5.1	8.1
Score 10 to 11	3.8	3.2
Score 12 to 15	6.1	7.1
Score 16 or Higher	5.7	7.6
WHODAS Total Score[2]		
Total Score (Range 0-24)[1]	3.3	3.4
Alternative Total Score (Range 0-8)[1]	0.8	0.9
MDE[3]		
Lifetime MDE	12.6	12.7
Past Year MDE	6.3	6.8
MDE TREATMENT[4]		
Saw or Talk to MD/Professional for MDE in Past Year	67.4[a]	*
Used Rx Medication for MDE in Past Year	59.5	53.1
MENTAL HEALTH TREATMENT		
Received Any Mental Health Treatment in Past Year[5]	13.6	14.0
Stay over in Hospital for Mental Health Treatment in Past 12 Months[5]	0.4[b]	1.2
Received Outpatient Mental Health Treatment in Past 12 Months[5]	6.9	6.4
Needed Mental Health Treatment but Didn't Get It in Past 12 Months	4.5	6.2
SUICIDAL THOUGHTS[6]		
Seriously Think about Killing Self in Past 12 Months	3.5	3.9
Make Plans to Kill Yourself in Past 12 Months	0.9	1.2
Try to Kill Yourself in Past 12 Months	0.2[b]	1.0
SUICIDAL THOUGHTS TREATMENT[6]		
Received Medical Attention Because Tried to Kill Self in Past 12 Months	0.1[b]	0.7
Stay Overnight at Hospital Because Tried to Kill Self in Past 12 Months	0.0[b]	0.6

K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; MHHS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

* Low precision.

NOTE: Table subset to only those cases who initially agreed to participate in the clinical follow-up interview. K6SCMAX is used in this table, which is defined as the higher K6 score between the past month K6 total score and the K6 total score in the worst month of the past year if the worst month was not the past 30 days.

[a] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.10$.
[b] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.05$.
[c] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.01$.
[1] Estimates are weighted means.
[2] For 2008, only sample A data are included.
[3] Respondents with unknown lifetime or past year MDE data were excluded.
[4] Among those with past year MDE.
[5] Mental health treatment/counseling is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.
[6] Respondents with unknown suicide information were excluded.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table D3c Substance Use Measures, by Participation in Clinical Interview: Persons Initially Agreeing to Participate in Clinical Follow-Up, 2008-2011 MHSS

Variable	Respondents to Clinical Follow-Up Estimate (%)	Final Nonrespondents Estimate (%)
LIFETIME		
ILLICIT DRUGS[1]	53.0	60.4
Marijuana and Hashish	49.5	54.7
Cocaine	16.8	20.0
Crack	2.9[b]	6.0
Heroin	1.6[a]	4.6
Hallucinogens	17.2	21.1
LSD	11.2	14.4
PCP	2.5	2.7
Ecstasy	5.5[a]	9.0
Inhalants	10.1	9.7
Nonmedical Use of Psychotherapeutics[2,3]	20.7	27.9
Pain Relievers	14.6	20.6
OxyContin®	2.0[b]	4.8
Tranquilizers	10.0	12.5
Stimulants[3]	9.6	9.3
Methamphetamine[3]	6.2	6.8
Sedatives	4.3[a]	2.7
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	31.0	38.1
ALCOHOL	89.6	86.2
TOBACCO[4]	72.5	78.8
CIGARETTES	67.2	*
PAST YEAR		
ILLICIT DRUGS[1]	15.5	19.1
Marijuana and Hashish	12.2	14.4
Cocaine	2.1	2.6
Crack	0.2[c]	0.8
Heroin	0.1[b]	0.6
Hallucinogens	1.5	1.5
LSD	0.2	0.4
PCP	0.0	0.1
Ecstasy	1.1	1.1
Inhalants	0.3	0.4
Nonmedical Use of Psychotherapeutics[2,3]	6.9	9.9
Pain Relievers	4.2	7.5
OxyContin®	0.4[b]	1.3
Tranquilizers	3.1	4.0
Stimulants[3]	2.3	1.9
Methamphetamine[3]	*	0.9
Sedatives	0.3	0.5
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	9.1	11.2
ALCOHOL	73.2[a]	61.8
TOBACCO[4]	33.2[b]	43.2
CIGARETTES	26.3[b]	39.3

(continued)

Table D3c Substance Use Measures, by Participation in Clinical Interview: Persons Initially Agreeing to Participate in Clinical Follow-Up, 2008-2011 MHSS (continued)

Variable	Respondents to Clinical Follow-Up Estimate (%)	Final Nonrespondents Estimate (%)
PAST YEAR SUBSTANCE ABUSE OR DEPENDENCE[5]		
ALCOHOL ABUSE OR DEPENDENCE	6.2[b]	12.4
PAIN RELIEVER ABUSE OR DEPENDENCE	0.5[b]	2.1

LSD = lysergic acid diethylamide; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; PCP = phencyclidine.

* Low precision.

NOTE: Table subset to only those cases who initially agreed to participate in the clinical follow-up interview.

[a] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.10$.
[b] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.05$.
[c] Difference between this estimate and corresponding estimate from final nonrespondents is statistically significant at $p < 0.01$.
[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. The estimates for Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine incorporated in these summary estimates do not include data from the methamphetamine items added in 2005 and 2006. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.
[2] Nonmedical use of prescription-type psychotherapeutics includes the nonmedical use of pain relievers, tranquilizers, stimulants, or sedatives and does not include over-the-counter drugs.
[3] Estimates of Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine in the designated rows include data from methamphetamine items added in 2005 and 2006 and are not comparable with estimates presented in NSDUH reports prior to the 2007 national findings report. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.
[4] Tobacco Products include cigarettes, smokeless tobacco (i.e., chewing tobacco or snuff), cigars, or pipe tobacco.
[5] Dependence or abuse is based on definitions found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Appendix E: Descriptive Characteristics for Overall Impact of Weighting Adjustments

Table E1 Descriptive Characteristics, by Sample and Response Status among Persons Aged 18 or Older, 2008-2011 Adult NSDUH Main Study and Clinical Sample

Variable	Full NSDUH Sample (1)	Selected MHSS Clinical Sample				
		Selected MHSS Clinical Sample (2)	MHSS Nonrespondents (3)	Respondents before Adjustment (4)	Respondents after Adjustment (Revised Weights) (5)	Respondents after Adjustment (Previous Weights) (6)
	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)
GENDER						
Male	48.3	48.2	52.8[b2]	45.1[a3]	48.3	48.3
Female	51.7	51.8	47.2[b2]	54.9[a3]	51.7	51.7
AGE						
18 to 25	14.8	16.2	15.6	16.6[a3]	14.8	14.8
26 to 34	15.8	14.4[a1]	12.7[b1a2]	15.5	17.2	15.5
35 to 49	27.5	28.1	32.9[b2]	24.9	26.1[a4]	27.8
50 or Older	41.9	41.3	38.8	43.0	41.9	41.9
AGE						
18 to 30	23.9	25.0	24.0	25.6	24.9	23.3
31 or Older	76.1	75.0	76.0	74.4	75.1	76.7
COUNTY TYPE						
Large Metro	53.2	56.4[a1]	59.1[a1]	54.6[a3]	50.8	52.0
Small Metro	30.6	28.3	26.2[b1]	29.7	31.8	30.7
Nonmetro	16.1	15.3	14.6	15.7	17.4	17.4
RACE/ETHNICITY						
White	68.0	66.4	59.9[b1b2]	70.8	68.0	68.0
Black	11.5	8.9[c1]	7.9[c1]	9.6[c3]	11.5	11.4
Hispanic	13.9	17.9[a1]	24.9[c1b2]	13.2	13.9	13.9
Asian and Others	6.6	6.8	7.3	6.4	6.6	6.7
EDUCATION						
Less than High School	15.1	14.5	22.3[a1c2]	9.3[c1b3]	12.1[a1c4]	8.2[c1]
High School Graduate	30.5	34.3[a1]	35.3	33.6[a3]	30.0	32.5
Some College	25.8	24.2	19.7[c1c2]	27.3	27.3	28.0
College Graduate	28.6	26.9	22.7[b1b2]	29.8	30.6	31.3
MARITAL STATUS						
Married	53.8	54.0	53.9	54.0	56.0	55.1
Widowed	6.0	4.5[b1]	5.2	4.0[b1]	3.4[c1]	3.4[c1]
Divorced or Separated	13.7	14.2	12.8	15.2	15.1	15.3
Never Married	26.5	27.3	28.1	26.8	25.6	26.2
INCOME						
Less than $20,000	18.2	16.0	18.2	14.5	13.8[c1a4]	12.2[c1]
$20,000 to $49,999	32.8	34.1	36.9	32.3	33.0	34.0
$50,000 to $74,999	17.4	17.7	13.4[c1b2]	20.6	19.6	19.6
$75,000 or More	31.6	32.2	31.4	32.7	33.5	34.2
REGION						
Northeast	18.5	21.8[b1]	20.4	22.7[a1]	22.5[a1]	21.3[a1]
Midwest	21.8	21.0	17.3[b1b2]	23.4	23.2	23.1
South	36.6	32.7[b1]	34.9	31.2[b1]	33.0[a1]	33.2
West	23.1	24.5	27.4	22.6	21.3	22.5
NOT CURRENTLY WORKING BUT WORKED IN THE PAST 12 MONTHS[1]						
Yes	7.8	7.1	4.9[c1b2]	8.6[a3]	7.3	7.7
Other[2]	92.2	92.9	95.1[c1b2]	91.4[a3]	92.7	92.3
RETIREMENT STATUS IN PAST WEEK						
No Job: Retired	14.9	13.3	15.8	11.6[b1]	12.6	12.0[a1]
Other[3]	85.1	86.7	84.2	88.4[b1]	87.4	88.0[a1]
WORK SITUATION IN PAST WEEK						
No Job: Looking for Work/ Layoff, Not Looking for Work/Disabled for Work	10.4	10.6	8.1[b1b2]	12.3	11.2	10.5
Other[4]	89.6	89.4	91.9[b1b2]	87.7	88.8	89.5

(continued)

Table E1 Descriptive Characteristics, by Sample and Response Status among Persons Aged 18 or Older, 2008-2011 Adult NSDUH Main Study and Clinical Sample (continued)

Variable	Full NSDUH Sample (1)	Selected MHSS Clinical Sample				
		Selected MHSS Clinical Sample (2)	MHSS Nonrespondents (3)	Respondents before Adjustment (4)	Respondents after Adjustment (Revised Weights) (5)	Respondents after Adjustment (Previous Weights) (6)
	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)
HEALTH INSURANCE[5]						
Covered by Any Health Insurance	82.1	82.4	77.6[b2]	85.7[c1]	85.3[c1]	84.9[a1]
Not Covered by Any Health Insurance	17.9	17.6	22.4[b2]	14.3[c1]	14.7[c1]	15.1[a1]
HAD DEPRESSION IN LIFETIME[6]						
Yes	12.6	13.0	10.3[a1b2]	14.8[b1a3]	13.4[c4]	15.3[c1]
No	87.4	87.0	89.7[a1b2]	85.2[b1a3]	86.6[c4]	84.7[c1]
HAD ULCER(S) IN LIFETIME[6]						
Yes	3.3	4.5	*	4.2	3.7[c4]	4.5
No	96.7	95.5	*	95.8	96.3[c4]	95.5
YEARS LIVED IN THE US						
Less than 5 years	1.7	1.4	2.0	1.0[b1]	1.0[b1]	0.9[c1]
At least 5 years but less than 10 years	2.2	2.5	3.9	1.5	1.3[a1]	1.1[c1]
10 years or more	11.7	13.8	17.8	11.2	10.3	9.1[a1]
Born in US	84.4	82.3	76.4[a1b2]	86.3	87.4[a1]	88.8[c1]
INTERVIEW PRIVACY						
Completely private	85.4	87.5[a1]	86.1	88.5[c1]	88.2[b1]	88.0[b1]
Minor distractions	10.6	9.3	10.7	8.3[c1]	8.3[c1]	8.6[c1]
Person(s) in the room or listening about 1/3 time	1.5	1.2	0.9[a1]	1.4	1.6	1.5
Serious interruptions of privacy more than 1/2 time	0.3	0.2	0.3	0.2	0.2	0.2
Constant presence of other person(s)	2.2	1.8	2.1	1.6	1.7	1.8
INTERVIEW LENGTH: Adult Depression						
Less than 2 minutes[7]	81.4	80.7	83.6	78.7[a1b3]	80.8[c4]	77.9[b1]
Greater or equal to 2 minutes	18.6	19.3	16.4	21.3[a1b3]	19.2[c4]	22.1[b1]
INTERVIEW LENGTH: Entire Interview						
Less than 60 minutes[7]	44.0	32.7[c1]	40.2[c2]	27.7[c1b3]	32.4[c1c4]	25.6[c1]
Greater or equal to 60 minutes	56.0	67.3[c1]	59.8[c2]	72.3[c1b3]	67.6[c1c4]	74.4[c1]

(continued)

Table E1 Descriptive Characteristics, by Sample and Response Status among Persons Aged 18 or Older, 2008-2011 Adult NSDUH Main Study and Clinical Sample (continued)

Variable	Full NSDUH Sample (1)	Selected MHSS Clinical Sample				
		Selected MHSS Clinical Sample (2)	MHSS Nonrespondents (3)	Respondents before Adjustment (4)	Respondents after Adjustment (Revised Weights) (5)	Respondents after Adjustment (Previous Weights) (6)
	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)
NO DIFFICULTY UNDERSTANDING THE MAIN NSDUH INTERVIEW						
No difficulty	90.5	89.8	84.0[a1b2]	93.7[b1]	92.2	93.6[b1]
Other[8]	9.5	10.2	16.0[a1b2]	6.3[b1]	7.8	6.4[b1]

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

* Low precision.

NOTE: For 2008 estimates (column 1), the 10 cases that were in neither sample (A or B) have been excluded; Column 1 weight= ANALWT_A, where ANALWT_A=ANALWT for 2009-2012, ANALWT_A=MHSAMPWT for 2008; Columns 2-4 weight= MHDSNWT, design-based weight; column 5 weight= MHNEWWGT, MHSS final revised weight; Column 6 weight= MHWEIGHT for 2008-2011, and MHSSWT1 for 2008 Sample A, previously adjusted weight.

[a1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.10$.

[b1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.05$.

[c1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.01$.

[a2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.10$.

[b2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.05$.

[c2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.01$.

[a3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.10$.

[b3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.05$.

[c3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.01$.

[a4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.10$.

[b4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.05$.

[c4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.01$.

[1] This question only includes respondents who said DK/REF to having worked in the last week or NO/DK/REF to DID YOU HAVE JOB BUSINESS in the past week.

[2] Other includes legitimate skip, did not work in the past 12 months, or missing.

[3] Other includes other no job categories, has job, or missing.

[4] Other includes other no job categories, has job, or missing.

[5] Respondents with unknown health insurance status were excluded.

[6] Respondents with unknown health data were excluded.

[7] Missing or negative data are included in this category.

[8] Other includes little difficulty, some difficulty, a lot of difficulty, or missing.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table E2 Mental Health, by Sample and Response Status among Persons Aged 18 or Older, 2008-2011 Adult NSDUH Main Study and Clinical Sample

Variable	Full NSDUH Sample (1)	Selected MHSS Clinical Sample (2)	Selected MHSS Clinical Sample: MHSS Nonrespondents (3)	Selected MHSS Clinical Sample: Respondents before Adjustment (4)	Selected MHSS Clinical Sample: Respondents after Adjustment (Revised Weights) (5)	Selected MHSS Clinical Sample: Respondents after Adjustment (Previous Weights) (6)
	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)
K6 Total Score						
K6 Total Score: Maximum of Past Year and Past 30 Days (Range: 0-24)[1]	4.9	4.8	4.2[b1c2]	5.2	5.1	5.3[b1]
Score 0 to 3	53.4	55.0	60.8[b1b2]	51.2	51.1	51.5
Score 4 to 5	13.7	12.8	11.3	13.8	13.4	13.0
Score 6 to 7	9.3	9.7	7.0[b1c2]	11.5[a1]	12.1[b1c4]	10.1
Score 8 to 9	6.1	5.9	6.6	5.4	5.6	6.0
Score 10 to 11	4.5	3.8[b1]	3.0[c1b2]	4.3	4.0	4.9
Score 12 to 15	6.7	6.5	5.7[a2]	7.0	7.6	7.1
Score 16 or higher	6.2	6.3	5.6	6.8[b3]	6.1[c4]	7.4
WHODAS Total Score[2]						
Total Score (Range 0-24)[1]	3.5	3.3	3.0[a1a2]	3.6	3.4[b4]	3.6
Alternative Total Score (Range 0-8)[1]	0.9	0.8	0.8	0.9	0.9	0.9
MDE[3]						
Lifetime MDE	12.8	12.6	9.5[c1c2]	14.7[b3]	12.9[c4]	14.6[a1]
Past Year MDE	6.6	6.4	5.3[b1b2]	7.2	6.6[a4]	7.2
MDE TREATMENT[4]						
Saw or Talk to MD/Professional for MDE in Past Year	61.8	64.5	56.3	68.5[b1a3]	65.9	66.6
Used Rx Medication for MDE in Past Year	51.7	57.7[b1]	53.4	59.7[b1]	56.4	56.9
MENTAL HEALTH TREATMENT						
Received Any Mental Health Treatment in Past Year[5]	13.6	13.7	12.2	14.7	13.7	14.2
Stay over in Hospital for Mental Health Treatment in Past 12 Months[5]	0.8	0.5[a1]	0.6	0.5	0.5	0.4[a1]
Received Outpatient Mental Health Treatment in Past 12 Months[5]	6.6	7.1	6.2	7.7[a1b3]	6.9[a4]	7.9[a1]
Needed Mental Health Treatment but Didn't Get It in Past 12 Months	4.9	4.8	4.2	5.2	4.8	5.4
SUICIDAL THOUGHTS[6]						
Seriously Think about Killing Self in Past 12 Months	3.7	3.5	2.8[c1b2]	4.1	3.7	3.8
Make Plans to Kill Yourself in Past 12 Months	1.0	1.0	0.8	1.0	1.1	1.0
Try to Kill Yourself in Past 12 Months	0.5	0.4	0.7[b2]	0.2[c1]	0.2[c1]	0.2[c1]

(continued)

Table E2 Mental Health, by Sample and Response Status among Persons Aged 18 or Older, 2008-2011 Adult NSDUH Main Study and Clinical Sample (continued)

Variable	Full NSDUH Sample (1)	MHSS Clinical Sample (2)	Selected MHSS Clinical Sample			
			MHSS Nonrespondents (3)	Respondents before Adjustment (4)	Respondents after Adjustment (Revised Weights) (5)	Respondents after Adjustment (Previous Weights) (6)
	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)
SUICIDAL THOUGHTS TREATMENT6						
Received Medical Attention Because Tried to Kill Self in Past 12 Months	0.3	0.2	0.4[b2]	0.1[c1]	0.1[c1]	0.1[c1]
Stay Overnight at Hospital Because Tried to Kill Self in Past 12 Months	0.2	0.2	0.4[b2]	0.0[c1]	0.0[c1]	0.0[c1]

K6 = Kessler-6, a 6-item psychological distress scale; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; WHODAS = 8-item World Health Organization Disability Assessment Schedule.

* Low precision.

NOTE: For 2008 estimates (column 1), the 10 cases that were in neither sample (A or B) have been excluded; Column 1 weight= ANALWT_A, where ANALWT_A=ANALWT for 2009-2012, ANALWT_A=MHSAMPWT for 2008; Columns 2-4 weight= MHDSNWT, design-based weight; column 5 weight= MHNEWWGT, MHSS final revised weight; Column 6 weight= MHWEIGHT for 2008-2011, and MHSSWT1 for 2008 Sample A, previously adjusted weight. K6SCMAX is used in this table, which is defined as the higher K6 score between the past month K6 total score and the K6 total score in the worst month of the past year if the worst month was not the past 30 days.

[a1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.10$.

[b1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.05$.

[c1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.01$.

[a2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.10$.

[b2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.05$.

[c2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.01$.

[a3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.10$.

[b3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.05$.

[c3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.01$.

[a4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.10$.

[b4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.05$.

[c4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.01$.

[1] Estimates are weighted means.

[2] For 2008, only sample A data are included.

[3] Respondents with unknown lifetime or past year MDE data were excluded.

[4] Among those with past year MDE.

[5] Mental health treatment/counseling is defined as having received inpatient care or outpatient care or having used prescription medication for problems with emotions, nerves, or mental health. Respondents were not to include treatment for drug or alcohol use. Respondents with unknown treatment/counseling information were excluded.

[6] Respondents with unknown suicide information were excluded.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.

Table E3 Substance Use Measures, by Sample and Response Status among Persons Aged 18 or Older, 2008-2011 Adult NSDUH Main Study and Clinical Sample

Variable	Full NSDUH Sample (1)	Selected MHSS Clinical Sample (2)	Selected MHSS Clinical Sample			
			MHSS Nonrespondents (3)	Respondents before Adjustment (4)	Respondents after Adjustment (Revised Weights) (5)	Respondents after Adjustment (Previous Weights) (6)
	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)
LIFETIME						
ILLICIT DRUGS[1]	49.5	52.9[a1]	50.3	54.7[b1]	54.7[b1]	55.6[c1]
Marijuana and Hashish	44.4	49.0[b1]	45.8	51.2[c1]	50.5[c1]	52.1[c1]
Cocaine	16.0	17.5	15.1	19.2	17.5	18.2
Crack	3.7	3.5	3.7	3.4	3.4	3.6
Heroin	1.7	2.3	2.8	1.9	1.9	2.2
Hallucinogens	15.7	17.6	14.0[b2]	20.0[a1]	18.0	18.8[b1]
LSD	10.1	11.9	9.7	13.4	11.3	12.5
PCP	2.7	2.6	1.9[a1]	3.0	2.6	3.0
Ecstasy	6.1	6.3	5.7	6.7[a3]	5.6[a4]	6.2
Inhalants	8.6	9.5	7.6[a2]	10.8[a1]	10.3	10.3
Nonmedical Use of Psychotherapeutics[2,3]	21.5	22.4	20.7	23.5	21.5	21.9
Pain Relievers	14.3	16.0	13.5	17.7	14.9	16.2
OxyContin®	2.3	2.8	3.4	2.4	2.1	2.0
Tranquilizers	9.2	11.8[a1]	10.5	12.6[b3]	8.8[c4]	11.2
Stimulants[3]	9.1	9.5	6.3[c1b2]	11.7	9.7	11.0
Methamphetamine[3]	5.5	6.3	4.1[b1]	7.8	6.0	7.2
Sedatives	3.5	3.6	2.1[c1c2]	4.7[a1]	3.8	4.3
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	31.4	32.6	30.7	34.0	32.6	33.0
ALCOHOL	87.4	86.3	81.8[a1a2]	89.4	89.8[a4]	91.8[c1]
TOBACCO[4]	73.4	71.5	69.0	73.1	73.9	75.2
CIGARETTES	68.8	65.8	62.1[a1]	68.3	68.4	70.3
PAST YEAR						
ILLICIT DRUGS[1]	14.4	16.4	13.3[a2]	18.4	16.3	16.9
Marijuana and Hashish	10.9	12.6	9.4[a2]	14.8	12.3	13.5
Cocaine	1.9	2.1	1.6[a2]	2.5	2.0	2.2
Crack	0.4	0.4	0.4	0.3	0.3	0.3
Heroin	0.2	0.3	0.4[a2]	0.2	0.2[a4]	0.2
Hallucinogens	1.5	1.5	1.2[a1]	1.8	1.8	1.6
LSD	0.3	0.3	0.2	0.3	0.3	0.4
PCP	0.0	0.0	0.0	0.0	*	*
Ecstasy	0.9	1.1	0.8	1.3	1.2	1.1
Inhalants	0.5	0.3[a1]	0.3[b1]	0.3	0.4	0.3
Nonmedical Use of Psychotherapeutics[2,3]	6.0	8.1	6.9	8.8	6.7	7.4
Pain Relievers	4.5	5.0	5.3	4.8	5.0	4.4
OxyContin®	0.6	0.7	1.0[a2]	0.5	0.5	0.4[b1]
Tranquilizers	2.1	3.9	3.0	*	2.2	3.5
Stimulants[3]	1.1	2.6	1.0	*	1.5	*
Methamphetamine[3]	0.4	*	0.5	*	0.9	*
Sedatives	0.3	0.4	0.3	0.4[a3]	0.3[b4]	0.4
ILLICIT DRUGS OTHER THAN MARIJUANA[1]	7.6	10.1	8.1	11.5	9.0	10.0
ALCOHOL	70.3	71.2	67.1	73.9	73.2	75.1[c1]
TOBACCO[4]	34.5	34.4	34.3	34.4	34.2	34.6
CIGARETTES	28.6	29.2	30.6	28.3	28.5	28.1

(continued)

Table E3 Substance Use Measures, by Sample and Response Status among Persons Aged 18 or Older, 2008-2011 Adult NSDUH Main Study and Clinical Sample (continued)

Variable	Full NSDUH Sample (1)	Selected MHSS Clinical Sample (2)	Selected MHSS Clinical Sample			
			MHSS Nonrespondents (3)	Respondents before Adjustment (4)	Respondents after Adjustment (Revised Weights) (5)	Respondents after Adjustment (Previous Weights) (6)
	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)	Estimate (%)
PAST YEAR SUBSTANCE ABUSE OR DEPENDENCE[5]						
ALCOHOL ABUSE OR DEPENDENCE	7.4	7.5	8.3	7.0	7.3	6.9
PAIN RELIEVER ABUSE OR DEPENDENCE	0.7	1.0	1.5[a1b2]	0.7	0.7[a4]	0.6

LSD = lysergic acid diethylamide; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; PCP = phencyclidine.

* Low precision.

NOTE: For 2008 estimates (column 1), the 10 cases that were in neither sample (A or B) have been excluded; Column 1 weight= ANALWT_A, where ANALWT_A=ANALWT for 2009-2012, ANALWT_A=MHSAMPWT for 2008; Columns 2-4 weight= MHDSNWT, design-based weight; column 5 weight= MHNEWWGT, MHSS final revised weight; Column 6 weight= MHWEIGHT for 2008-2011, and MHSSWT1 for 2008 Sample A, previously adjusted weight.

[a1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.10$.

[b1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.05$.

[c1] Difference between this estimate (from columns 2-6) and corresponding estimate from full CAI sample (column 1) is statistically significant at $p < 0.01$.

[a2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.10$.

[b2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.05$.

[c2] Difference between this estimate from nonrespondents (column 3) and corresponding estimate from respondents (column 4) using design-based weights is statistically significant at $p < 0.01$.

[a3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.10$.

[b3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.05$.

[c3] Difference between this estimate (column 4) using design-based weights and corresponding estimate from respondents (column 5) using final adjusted weights is statistically significant at $p < 0.01$.

[a4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.10$.

[b4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.05$.

[c4] Difference between this estimate (column 5) using final adjusted weights and corresponding estimate from respondents (column 6) using initial adjusted weights is statistically significant at $p < 0.01$.

[1] Illicit Drugs include marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. Illicit Drugs Other than Marijuana include cocaine (including crack), heroin, hallucinogens, inhalants, or prescription-type psychotherapeutics used nonmedically. The estimates for Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine incorporated in these summary estimates do not include data from the methamphetamine items added in 2005 and 2006. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.

[2] Nonmedical use of prescription-type psychotherapeutics includes the nonmedical use of pain relievers, tranquilizers, stimulants, or sedatives and does not include over-the-counter drugs.

[3] Estimates of Nonmedical Use of Psychotherapeutics, Stimulants, and Methamphetamine in the designated rows include data from methamphetamine items added in 2005 and 2006 and are not comparable with estimates presented in NSDUH reports prior to the 2007 national findings report. See Section B.4.8 in Appendix B of the *Results from the 2008 National Survey on Drug Use and Health: National Findings*.

[4] Tobacco Products include cigarettes, smokeless tobacco (i.e., chewing tobacco or snuff), cigars, or pipe tobacco.

[5] Dependence or abuse is based on definitions found in the 4th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, NSDUH main study and clinical sample, 2008-2012.