2012 National Survey on Drug Use and Health

ESTIMATING MENTAL ILLNESS AMONG ADULTS IN THE UNITED STATES: REVISIONS TO 2008 ESTIMATION PROCEDURES

DISCLAIMER

SAMHSA provides links to other Internet sites as a service to its users and is not responsible for the availability or content of these external sites. SAMHSA, its employees, and contractors do not endorse, warrant, or guarantee the products, services, or information described or offered at these other Internet sites. Any reference to a commercial product, process, or service is not an endorsement or recommendation by SAMHSA, its employees, or contractors. For documents available from this server, the U.S. Government does not warrant or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed.

Substance Abuse and Mental Health Services Administration
Rockville, Maryland 20857

September 2015

This page intentionally left blank

2012 NATIONAL SURVEY ON DRUG USE AND HEALTH: ESTIMATING MENTAL ILLNESS AMONG ADULTS IN THE UNITED STATES: REVISIONS TO 2008 ESTIMATION PROCEDURES

Contract No. HHSS283201000003C
RTI Project No. 0212800.001.112.008.009.004

Authors:

Phillip Kott
Sarra Hedden
Jeremy Aldworth
Jonaki Bose
James Chromy
Joseph Gfroerer
Dan Liao

Project Director:

David C. Hunter

SAMHSA Project Officer:

Peter Tice

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

September 2015

Recommended Citation: Center for Behavioral Health Statistics and Quality. (2015). *Estimating Mental Illness Among Adults in the United States: Revisions to the 2008 Estimation Procedures*. Substance Abuse and Mental Health Services Administration, Rockville, MD.

Acknowledgments

This report was prepared by the Center for Behavioral Health Statistics and Quality (CBHSQ), Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services (HHS), and by RTI International (a trade name of Research Triangle Institute), Research Triangle Park, North Carolina. Work by RTI was performed under Contract No. HHSS283201000003C.

Table of Contents

Chapter **Page**

1. Introduction 1
   - 1.1 The 2008 (WHODAS) Estimation Methods 3
   - 1.2 Estimating Other Categories of Mental Illness 3
   - 1.3 Evaluating and Replacing the 2008 Estimation Methods 4

2. Statistical Underpinnings 9
   - 2.1 The Standard Cut Point and Probability Estimators 9
   - 2.2 Bias Measures at the Domain Level 12
   - 2.3 Bias-Corrected Estimators 13
   - 2.4 Variances and Mean Squared Errors 13
     - 2.4.1 All Adults 13
     - 2.4.2 Domain Level 14
   - 2.5 Evaluating the Models Underlying a Standard Cut Point Estimator 16
   - 2.6 Simultaneously Estimating SMI and AMI with a Cut Point Estimator 17
   - 2.7 Weighting the Clinical Subsample Data 19

3. Choosing an Estimator 21

4. Evaluating Alternative Models 27
   - 4.1 The Search for New Models 27
     - 4.1.1 Predictor Variables 27
     - 4.1.2 Modeling Steps 30
   - 4.2 Comparing Models 31
   - 4.3 Removing Bias at the Domain Level 33

5. The 2012 Models and Estimators 37
   - 5.1 The 2012 WHODAS Estimation Method 37
   - 5.2 The 2012 Model for the 2008 SDS Sample 44

6. Comparisons of SMI and AMI Prevalence Estimates Using the 2012 and 2008 Models 47

7. Some Concluding Comments 53
   - 7.1 The Target Population 53
   - 7.2 Comparability of NSDUH's SMI and AMI Prevalence Estimates with Other National Estimates 53
   - 7.3 The Sensitivity and Specificity of SMI and AMI Prevalence Estimates 56
   - 7.4 Limitations in Using the Mental Illness Variables in Analyses 57

References 63

Appendix

A Comparisons of Estimates Using the 2008 and 2012 Models A-1

List of Tables

Table **Page**

1.1 Mental Illness Categories Defined by SCID Disorder Diagnosis and GAF Score: 2008-2012 MHSS Clinical Follow-Up Study 4

1.2 Overview of Investigations and the Report 6

2.1 Logistic Model Coefficients with Different Dependent Variables 18

3.1 The Estimators for SMI among All Adults Using the 2012 (WHODAS) Model 21

3.2 The Standard Cut Point and Probability Estimators for SMI and Their Bias Measures Using the 2012 Model by Domain 22

3.3 The Two Cut Point Estimators for SMI (Standard and Hybrid) and Their Bias Measures Using the 2012 Model by Domain 24

4.1 Evaluating the Standard Cut Point Estimators for All Adults Based on Alternative SMI Models 32

4.2 Evaluating the Standard Cut Point Estimators for All Adults Based on Alternative AMI Models 33

4.3 Some Standard Cut Point Estimators for SMI and Their Bias Measures Using a Model with Kc_7, Ac_0, MHSUITHK, and AMDEYR2 34

4.4 Evaluating SMI and AMI Cut Point Estimators for All Adults Employing Different Variables in the Model 35

5.1 Coefficients of Logistic Models for SMI and AMI with AGE1830 and AGE30 as Predictor Variables 38

5.2 Summary Output of Standard Cut Point Estimates Using the Model Predicting SMI (SMI Method) 41

5.3 Summary Output of Standard Cut Point Estimates Using the Model Predicting AMI (AMI Method) 42

5.4 Summary Output of Standard Cut Point Estimates Using a Multivariate Prediction Model (Combined Method) 43

5.5 Comparing Standard Cut Point Estimates for SMI from the 2008B Sample Using the 2012 Model for the SDS Sample with 2012 Method Estimates from the 2008A Sample by Domain 45

5.6 Comparing Standard Cut Point Estimates for AMI from 2008B Sample Using the 2012 Model for the SDS Sample with 2012 Method Estimates from the 2008A Sample by Domain 46

6.1 The 2008 and 2012 Models and Their Cut Points 48

6.2 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models by Domain 49

6.3 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models by Domain 51

6.4 Comparing Mental Health Prevalence Estimates Using the 2008 and 2012 Models, by Detailed Age Group 52

7.1 Sensitivity and Specificity of Alternative Models for SMI in Producing SMI and AMI All-Adult Prevalence Estimates 57

7.2 Comparing SMI and AMI Prevalence Estimates for MDE and Suicidal Thoughts Domains Using the 2008 and 2012 Models 58

7.3 Comparing SMI Prevalence Estimates for K6 and WHODAS Domains Using the 2008 and 2012 Models 60

7.4 Comparing AMI Prevalence Estimates for K6 and WHODAS Domains Using the 2008 and 2012 Models 61

7.5 Comparing Proportions Planning or Attempting Suicide in the Past Year Among Adults with SMI Using the 2008 and 2012 Models 62

7.6 Comparing Proportions Planning or Attempting Suicide in the Past Year Among Adults with Suicidal Thoughts by SMI Domain Using the 2008 and 2012 Models 62

A.1 Comparing Standard Cut Point Estimates for the SMI 2008B Sample Using a 2012 Model for that Sample Containing an SDS Variable in Place of Ac_0 with Estimates from the 2008A Sample by Domain A-1

A.2 Comparing Standard Cut Point Estimates for the AMI 2008B Sample Using a 2012 Model for that Sample Containing an SDS Variable in Place of Ac_0 with Estimates from the 2008A Sample by Domain A-2

A.3 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008A NSDUH by Domain A-3

A.4 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008B NSDUH by Domain A-4

A.5 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008 NSDUH by Domain A-5

A.6 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2009 NSDUH by Domain A-6

A.7 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2010 NSDUH by Domain A-7

A.8 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2011 NSDUH by Domain A-8

A.9 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2012 NSDUH by Domain A-9

A.10 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008A NSDUH by Domain A-10

A.11 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008B NSDUH by Domain A-11
A.12 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008 NSDUH by Domain A-12
A.13 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2009 NSDUH by Domain A-13
A.14 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2010 NSDUH by Domain A-14
A.15 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2011 NSDUH by Domain A-15
A.16 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2012 NSDUH by Domain A-16

1. Introduction

As part of the Mental Health Surveillance Study (MHSS), the Substance Abuse and Mental Health Services Administration (SAMHSA) conducted a clinical follow-up survey via telephone on a subsample of the National Survey on Drug Use and Health (NSDUH) adult respondents from 2008 to 2012 to assess the prevalence of selected mental disorders. The purpose of this clinical data collection was to use the data to develop statistical models that would provide national and state estimates of serious mental illness (SMI) when combined with NSDUH data.

Public Law No. 102-321, the Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA) Reorganization Act of 1992, established a block grant for states within the United 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 SAMHSA to develop a definition for the term "adults with SMI." SAMHSA defined adults with SMI as individuals aged 18 or older who currently or at any time in the past year have had a diagnosable mental, behavioral, or emotional disorder (excluding developmental and substance use disorders) of sufficient duration to meet diagnostic criteria specified within the *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition, Text Revision (DSM-IV-TR) (American Psychiatric Association [APA], 2000) that has resulted in serious functional impairment, which substantially interferes with or limits one or more major life activities. For more details, see Section B.4.3 in Appendix B of the 2012 mental health findings report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2013). The MHSS clinical data were used in developing a statistical model to apply to the full NSDUH sample that would generate estimates of the annual percentage of adults aged 18 or older who have SMI at national and state levels.

In NSDUH reports prior to 2004, the Kessler 6 (K6) psychological distress scale was used to measure SMI (from 2001[1] to 2003). In 2004, yearly estimation of SMI ceased temporarily because of concerns about the validity of using only the K6 distress scale to measure SMI without including a functional impairment scale (see Section B.4.4 of Appendix B in the 2004 NSDUH national findings report [Office of Applied Studies, 2005] for a discussion). In December 2006, a technical advisory group was convened by SAMHSA's Office of Applied Studies (OAS, which later became CBHSQ) and the Center for Mental Health Services (CMHS) to solicit recommendations for mental health surveillance data collection strategies to address SAMHSA's legislative requirements. The panel recommended that NSDUH be used to produce estimates of SMI among adults.

Recognizing the limitations on conducting full clinical interview assessments of all NSDUH respondents, the panel suggested that the K6 psychological distress scale, already included in the NSDUH, be supplemented with questions on functional impairment, as both of these scales in combination were expected to be strong predictors of SMI. The panel also suggested that the data from these short scales be used to estimate SMI, using a statistical model based on clinical interviews conducted on a subsample of NSDUH respondents. As a result, the

[1] The survey was known as the National Household Survey on Drug Abuse in 2001.

approach adopted by 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 from the main NSDUH adult samples for each year from 2008 to 2012 were produced. The model continued to be used to produce yearly SMI estimates after the clinical study ended.

In 2008, additional questions assessing functional impairment were added to the NSDUH main-survey interview. Although all respondents received the K6 to assess psychological distress, a split-sample design was used to randomly assign one of two functional impairment scales, the abbreviated World Health Organization Disability Assessment Schedule (WHODAS; Rehm et al., 1999 and abbreviated in Novak, Colpe, Barker, & Gfroerer, 2010) or the Sheehan Disability Scale (SDS; Leon, Olfson, Portera, Farber, & Sheehan, 1997). Both functional impairment instruments were added to determine which one more accurately predicted SMI as measured via clinical interviews in the subsample of approximately 1,500 adult respondents from the NSDUH main survey (herein referred to as the "clinical subsample") who completed a follow-up clinical interview to assess presence or absence of mental illness and level of functional impairment to use as the gold standard clinical subsample.

Each participant in the 2008 MHSS clinical study was administered the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I).[2] This SCID was adapted by the developer to be administered over the telephone by masters- or doctoral-level clinical interviewers who had undergone extensive training with clinical supervisors and the developer of the SCID. The clinical interviews were administered approximately 2 to 4 weeks after the NSDUH interview. Functional impairment ratings in the SCID were assigned by clinical interviewers using the Global Assessment of Functioning (GAF) scale.[3] For the clinical interview data, persons were defined as having SMI if they had a diagnosable mental, behavioral, or emotional disorder in the past 12 months (other than a developmental or substance use disorder) that met DSM-IV criteria (APA, 1994) and resulted in substantial functional impairment. An evaluation was performed to determine which combination of K6 and impairment scale (WHODAS or SDS) collected in NSDUH worked best in the statistical model used to predict SMI status as measured in the MHSS clinical subsample.

Based on this evaluation it was determined that the WHODAS was a better predictor of SMI than the SDS and that the WHODAS would continue to be administered as the sole impairment scale in the 2009 and subsequent NSDUHs (Aldworth et al., 2010a). Therefore, the remainder of this section focuses on the model that was fit on the half sample that was assigned to the WHODAS in 2008. This model is called the "2008 (WHODAS) model." A similar model

[2] 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).

[3] 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 scores higher than 60 represent several categories of impairment ranging from none to slight, and scores lower than 51 represent several categories ranging from serious to extreme.

was constructed in the 2008 MHSS for the half sample that was assigned to the SDS. Section 6 provides more details on the SDS half sample.

1.1 The 2008 (WHODAS) Estimation Methods

As described previously, roughly half of the full adult NSDUH sample in 2008, denoted as the 2008A (half) sample, were assigned functional impairment questions derived from the WHODAS. (The half sample assigned to the SDS functional impairment questions is referred to as the 2008B half sample.) A weighted logistic regression model was fit on diagnostic data collected from a subsample of the 2008A sample who received a clinical interview in the MHSS (n = 759). The dependent variable for the model was SMI (1 = yes, 0 = no) as measured by the MHSS and the predictor (explanatory) variables were based on the psychological distress (K6) and functional impairment (WHODAS) items collected in the NSDUH main interview. Then, this 2008 (WHODAS) model was used to produce a predicted probability of having SMI for each clinical interview respondent in the 2008A half sample. A cut point was established among the fitted probabilities of having SMI 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 adults in the clinical subsample 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) within the clinical subsample of the 2008A sample.

Because the predictor variables in the model (i.e., K6 and WHODAS variables) were collected in the NSDUH 2008A half sample, a probability of having SMI can be computed for every NSDUH adult respondent in the 2008A half sample using the estimated model parameters. Using the predicted probabilities and the cut point derived from the clinical subsample, the dichotomized SMI predicted values were computed for the half sample. 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 respondents' predicted probabilities. The SMI predicted values for the 2008A half sample then were weighted appropriately to compute prevalence estimates of SMI for all eligible adults in 2008. Moreover, SMI probabilities and SMI predicted values were computed for all adult respondents in the NSDUH full samples from 2009 to 2011 using the same model parameters based on the 2008A clinical subsample. More details on the 2008 model and the mental illness prevalence estimates it produces are contained in Chapter 6.

1.2 Estimating Other Categories of Mental Illness

A primary objective of the MHSS clinical subsample was to produce annual national estimates of SMI prevalence that were accurate for all adults and for adult subpopulations. Another objective was to produce estimates of other mental illness categories defined by level of impairment, such as mild (or low) mental illness (LMI), moderate mental illness (MMI), and, more critically, any mental illness (AMI).

These other categories of mental illness are based on SCID disorder diagnoses and GAF scores as defined in Table 1.1. A respondent to the MHSS clinical subsample 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 (the higher the GAF score the more functional adept the respondent). AMI, defined

as having a mental disorder regardless of the level of impairment because of that disorder, is the category obtained by collapsing the first three categories in Table 1.1 into a single category. This report focuses primarily on SMI, to a lesser extent on AMI, and occasionally on the combination category serious or moderate mental illness (SMMI).

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 ≤ 50 (severe or worse impairment)
Moderate Mental Illness (MMI)	One or More	50 < GAF ≤ 59 (moderate impairment)[1]
Mild 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; 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 ≤ 60. The cutoff of 59 for MMI and LMI was chosen to conform to the corresponding cutoff selected by Kessler and colleagues (2003).

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

The probabilities of having SMI derived from the 2008 MHSS regression model (i.e., 2008 WHODAS model) also were used to produce AMI predicted values for adult respondents in the NSDUH full sample in the following manner. A second cut point was determined so that if adults with probabilities at or above the cut point were predicted to have AMI and the rest were not, the weighted total of false AMI positives and false AMI negatives in the clinical subsample would come as close as possible to being equal.

Although SAMHSA continued to obtain clinical interviews after 2008, estimates of SMI and AMI from the 2009, 2010, and 2011 NSDUHs have been based on the clinical subsample for the 2008 WHODAS model. The same model was applied to each year's NSDUH data to provide comparability in estimating SMI and AMI trends across the years. Producing a new model each year based on the small annual clinical subsamples (only 500 interviews in 2009 and 2010) would have resulted in changes in the model parameters and corresponding prevalence rates solely because of sampling error, making it difficult to detect real trends in SMI and AMI over time. Furthermore, an evaluation of the 2008 model using 2009 clinical data found that the model could not be significantly improved with the additional 500-case 2009 clinical subsample (Aldworth et al., 2010b).

1.3 Evaluating and Replacing the 2008 Estimation Methods

With financial support from the National Institute of Mental Health (NIMH) through an interagency agreement, SAMHSA expanded the clinical interview subsample in 2011 and 2012 and began several studies to investigate the accuracy of the estimates produced using the 2008 model. By the end of 2012, a nationally representative sample of more than 5,500 completed clinical assessments had been collected (including those assigned SDS questions in the 2008 main interview). This is the largest nationally representative sample of clinical diagnostic interviews of mental disorders by clinically trained interviewers known to exist in the United States. This larger dataset was used to evaluate the accuracy of SMI and AMI estimates based on the 2008 model and estimation method (including how the sample should be weighted). If a

revised model and weights could be shown to produce more accurate estimates, improved annual mental illness prevalence estimates for 2012 would be published. In addition, the previously released (2008 to 2011) estimates of SMI and AMI would be revised so that trends across survey years could be better measured and more accurate portrayals of SMI and AMI produced.

A previous study suggested that refitting the 2008 model with both 2008 and 2009 data would not be prudent since changes in the estimates would reflect little more than random noise (Aldworth et al., 2010b). Nevertheless, accuracy gains were more likely to be realized from fitting the same model to a much larger dataset—one that by the end of 2012 included the 759 respondents from 2008, 520 respondents from 2009, 516 from 2010, 1,495 from 2011, and 1,622 from 2012. Moreover, the possibility of improving the estimation methods for SMI prediction needed to be explored.

A parallel investigation to the research described in this report focused on how the clinical subsample data should be weighted both for modeling and for direct estimation (CBHSQ, 2014). A brief review of the weighting changes implemented to reduce the impact of undercoverage and nonresponse in the clinical subsample is discussed in Chapter 2 (Section 2.7).

The new weighting scheme recognized that Hispanics who responded to the NSDUH main-survey interview in Spanish were ineligible for the clinical subsample and adjusted the weights of Hispanics with otherwise similar characteristics (e.g., age, education) to compensate. There were also separate nonresponse adjustments for adults selected for the clinical subsample who refused to be recontacted and those who agreed to be recontacted but were not, which could happen for a variety of reasons. These adjustments better reflected and compensated for the causes of nonresponse. It appears that, among adults selected for the clinical interview, the ones in the SMI clinical subsample were more likely to complete the clinical interview than were the ones with similar characteristics who did not have SMI.

In addition to changes in how the clinical subsamples were weighted each year, Section 2.7 discusses how the clinical subsamples were combined across years for modeling purposes under the assumption (supported by statistical testing) that the underlying relationship between SMI and its covariates did not change across the years from 2008 to 2012. Most of the analyses in this report use weights combined in this manner except when stated otherwise. Major exceptions are model-derived estimates in the Chapter 6 tables since they display what would actually be produced using the alternative models and NSDUH main-study weights.

The preliminary statistics in Chapter 2 introduce alternative estimators and the criteria both for choosing among the estimators and for choosing among different SMI and AMI models for a given estimator. The subsequent chapters address these questions:

1. Would employing a different estimator for SMI prevalence substantially improve the resulting prevalence estimates (Chapter 3)?
2. If no substantial improvement on the SMI prevalence estimator is available, is it possible to improve the model used substantially with the original cut point estimator, either by re-estimating the model parameters using the larger sample or by including different predictor variables (Chapter 4)?

3. Would there be substantive improvements in the estimation of AMI without harm to the estimation of SMI if cut point were based on the model for AMI rather than for SMI? What about a joint model for SMI and AMI (Chapter 5)?

An additional question prompted by changes in the estimation method using the NSDUH samples from 2008 through 2012 is whether the 2008B data (based on a model using SDS rather than WHODAS) could be incorporated into the estimation (Section 5.2).

Chapter 5 finalizes the reasons for adapting what is called the "2012 WHODAS model for SMI" along with a (standard) cut point estimator when estimating mental-illness prevalences (SMI and AMI). In addition, a modified version of the model is proposed and adapted for the 2008B SDS data (which has no WHODAS variables). In much of this report the 2012 WHODAS model for SMI is called simply the "2012 model."

Chapter 6 compares estimates for SMI and AMI using the 2008 and 2012 models. Chapter 7 provides some concluding comments about the limitations of NSDUH-based SMI and AMI prevalence estimates. Table 1.2 provides an overview the investigations and the report.

Table 1.2 Overview of Investigations and the Report

Chapter	Issues	Investigations	Decisions or Findings
3	Which of three estimators is best for estimating SMI prevalence?	Compare three estimators of total adult prevalence to direct estimator in terms of prevalence, bias, and RMSE. Compare prevalence estimates, bias, and SE of bias for 32 domains.	Use standard cut point estimator.
4	Which covariates should be considered for a new model using multiple years of data (2008-2011)?	• Summarize possible predictor variables. • Construct models for SMI and AMI. • Remove from consideration any model that does not predict monotonically with certain predictor variables. • Compute error rates and proxy SEs for total adults. • Narrow the set of predictor variables to consider. • Adjust the models to reduce bias for domains. • Summarize recommendations for covariates.	• For SMI and AMI of total adults, consider models that use these variables: • Kc_7 • Ac_0 • MSSUITHK • AMDEYR2 • AMDELT2 • To reduce bias in domain estimates, use an age variable such as AGE1837 or AGE1830. • See Chapter 4 for variable definitions.

(continued)

Table 1.2 Overview of Investigations and the Report

Chapter	Issues	Investigations	Decisions or Findings
5	• Should MHSUITHK and AMDEYR2 be added to the model? • Which, if any, of three candidate age variables should be included? • Which of three methods should be used to ensure that SMI and AMI predictions are consistent for an individual respondent?	• Combine the data across years, scaled and unscaled. • Compare model parameter estimates when different forms of age variables are used. • Compare error rate and RMSE on prevalence estimates for all adults for various combinations of covariates.	• Note the basic model structure is stable across time. • Add MHSUITHK and AMDEYR2 to the WHODAS model, along with Kc_7 and AC_0. • Include an age variable, which may be helpful for estimating all adults, not just domains. • Add AGE1830 to the model to remove domain-level biases. • Use SMI cut point method for ensuring consistency.
5	Can the new 2012 model be used with the 2008 SDS half sample?	Compare prevalence estimates and bias estimates for the two 2008 half samples and the direct estimates.	• Prevalence estimates are very similar for the two half samples. • Both half samples are similar to the direct estimates. • Bias estimates were acceptable. • The 2012 model can be used with the 2008 SDS half sample.
6	What are the ramifications of using the 2012 model rather than the 2008 model for estimating multiyear prevalence?	• Compare model parameters and cut points. • Compare prevalence estimates for all adults from both models with direct estimates for domains. • Compare naïve SE estimates for the two models.	• 2012 model estimates were closer to direct estimates. • 2012 model estimates had less bias than 2008 model estimates, especially for the 18-25 age group. • Naïve SEs were generally smaller for the 2012 model.
7	What other considerations should be noted?	• Review target population. • Explore why NSDUH SMI estimates are smaller than those from other surveys. • Discuss how error rate relates to sensitivity and specificity. • Provide caution about use of models.	Several factors may contribute to survey differences: • Target population • Covered population • Methodology • Survey mode • Specific measures • Interviewer training • Context effects • Analyses involving variables use in SMI model restrictions can be biased.

AMI = any mental illness; RMSE = root mean square errors; SDS = Sheehan Disability Scale; SE = standard error; SMI = serious mental illness.

This page intentionally left blank

2. Statistical Underpinnings

In this chapter, we first present the statistical theory behind three different estimators for serious mental illness (SMI) prevalence based on the same logistic model fit of SMI in the clinical subsample. The three estimators are the standard cut point, probability, and hybrid cut point estimators. Section 2.1 describes and provides mathematical expressions for these three estimators. Section 2.2 investigates the bias measure for each estimator at the domain level, while Section 2.3 proposes the creation of bias-corrected estimators. Section 2.4 discusses measuring the root mean square errors (RMSEs) of these three estimators.

Because the standard cut point estimator was ultimately used to produce both SMI and any mental illness (AMI) prevalence estimates, Section 2.5 introduces criteria for evaluating alternative models when a cut point estimator is to be used. Section 2.6 addresses different ways of making cut point estimates for SMI and AMI logically consistent (e.g., an adult National Survey on Drug Use and Health [NSDUH] respondent predicted to have SMI must also be predicted to have AMI).

Estimates computed from the NSDUH main survey and for the Mental Health Surveillance Study (MHSS) clinical subsample incorporate survey weights to compensate for unequal probabilities of selection, noncoverage and nonresponse biases. Developing weights that were determined for analytical purposes is briefly discussed in Section 2.7.

2.1 The Standard Cut Point and Probability Estimators

Let S denote the combined NSDUH World Health Organization Disability Assessment Schedule (WHODAS) half sample from 2008 combined with the full NSDUH samples from 2009 through 2012 (i.e., all adult respondents who were administered the WHODAS in the NSDUH main interview). Let S' be the WHODAS clinical subsample, with w_k denoting the NSDUH WHODAS full sample weight and ω_k denoting the WHODAS clinical subsample weight for a NSDUH respondent $k \in S$. For convenience, we set ω_k to 0 for a respondent in S but not S'. Let $y_k = 1$ when $k \in S'$ was diagnosed to have SMI, and $y_k = 0$ when $k \in S'$ was diagnosed not to have SMI.[4]

Under probability-sampling theory (also called design-based sampling theory), a nearly unbiased estimator for SMI prevalence among adults in this time period (2008 to 2012) is the *direct estimator* computed using only the clinical respondents:

[4] This report assumes being diagnosed in the clinical evaluation with a disorder (other than development or substance use disorder) and having psychological distress and functional impairment to the level of SMI as measured by the NSDUH questionnaire is equivalent to actually having SMI. Although not strictly accurate, we sometimes use the simpler language that the respondent was diagnosed with SMI. (We use similar language shortcuts for AMI and other categories of mental illness.) A weaker assumption is that for every domain of interest, the fraction of adults that would be misdiagnosed as having SMI using the Structured Clinical Interview for DSM-IV (SCID) is equal to the fraction that would be misdiagnosed as not having SMI.

$$\bar{y}_U = \frac{\sum_{k \in S'} \omega_k y_k}{\sum_{k \in S'} \omega_k}.$$

(Since the denominator is a random variable, a ratio estimator like $\bar{y}_U$ is only "nearly" unbiased).

Now, suppose a ω_k-weighted logistic regression is run on the clinical respondents in S' with y_k as the dependent variable and a vector of explanatory covariates (predictors) $\mathbf{x}_k$ that includes an intercept. Let the predicted variable for y_k from this weighted logistic regression be $p_k = p(\mathbf{x}_k'\mathbf{b}) = [1 + \exp(-\mathbf{x}_k'\mathbf{b})]^{-1}$, where $\mathbf{b}$ is the estimated parameter vector from the logistic regression, and where p_k is the resulting estimated probability that adult k has SMI.

Sorting the subsample S' by the p_k values, a cut point value p_C can be found such that

$$\sum_{\substack{k \in S' \\ p_k \geq p_C}} \omega_k = \sum_{k \in S'} \omega_k y_k \tag{1}$$

holds exactly or as closely as possible. That is to say, the estimated number of adults in the population having p_k-values at or above the cut point based on the MHSS clinical subsample approximately equals the number of adults diagnosed with SMI in the subsample. A cut point determined using equation (1) also comes as close as possible to equalizing the weighted total of false positives and false negatives in S'.

Every adult in S can be assigned p_k values based on $\mathbf{x}_k$ using $p_k = [1 + \exp(-\mathbf{x}_k'\mathbf{b})]^{-1}$. After defining the indicator random variable c_k to be 1 when $p_k \geq p_C$ and 0 otherwise, two alternative estimators for SMI prevalence among adults are the model-derived *standard cut point* ($\bar{y}_C$) and *probability* ($\bar{y}_P$) estimators:

$$\bar{y}_C = \frac{\sum_S w_k c_k}{\sum_S w_k}, \tag{2}$$

and

$$\bar{y}_P = \frac{\sum_S w_k p_k}{\sum_S w_k}. \tag{3}$$

Unlike $\bar{y}_U$, the standard cut point and probability estimates are calculated from the full NSDUH (adult) respondent sample, S. The former is the estimated fraction of adults assigned SMI status based on the full NSDUH. The latter is the estimated average probability an adult has SMI based on full NSDUH. Those two fractions are as close as possible to equal when based on the clinical subsample, but *not* on the full NSDUH sample.

To estimate SMI prevalence in a domain d (e.g., among males), an indicator for domain membership, d_k, can be inserted into the estimators for all adults:

$$\bar{y}_{U(d)} = \frac{\sum_{S'} \omega_k y_k d_k}{\sum_{S'} \omega_k d_k} \tag{4}$$

$$\bar{y}_{C(d)} = \frac{\sum_S w_k c_k d_k}{\sum_S w_k d_k} \tag{5}$$

$$\bar{y}_{P(d)} = \frac{\sum_S w_k p_k d_k}{\sum_S w_k d_k} \tag{6}$$

The indicator (d_k) is defined to be 1 when k is in the domain and 0 otherwise.

A hybrid of the probability and standard cut point estimators is also of interest. Suppose the full NSDUH respondent sample (and not just clinical subsample respondents) were sorted by their estimated p_k values. A hybrid cut point p_H can be determined in the NSDUH sample so that

$$\sum_{\substack{k \in S \\ p_k \geq p_H}} w_k = \sum_{k \in S} w_k p_k \tag{7}$$

holds as closely as possible. Setting $h_k = 1$ when $p_k > p_H$ and 0 otherwise. The hybrid cut point estimator for SMI prevalence in a domain is

$$\bar{y}_{H(d)} = \frac{\sum_S w_k h_k d_k}{\sum_S w_k d_k}. \tag{8}$$

By design, if a p_H could be found that satisfies equation (7) exactly, then the hybrid cut point estimator would exactly equal the probability estimator for the NSDUH adult full sample.

If no perfect hybrid cut points exist, then the hybrid cut point estimator for all adults would have a slight bias, which could be measured, squared, and then added to the variance of the probability estimator to compute its mean squared error.

2.2 Bias Measures at the Domain Level

The direct estimator in equation (4) is nearly unbiased in every domain. To measure the bias for the estimators in equations (5), (6), or (8), respectively, one can use the following estimators computed using only MHSS clinical subsample[5]:

$$B_{C(d)} = \frac{\sum_{S'} \omega_k (c_k - y_k) d_k}{\sum_{S'} \omega_k d_k}, \tag{9}$$

$$B_{P(d)} = \frac{\sum_{S'} \omega_k (p_k - y_k) d_k}{\sum_{S'} \omega_k d_k}, \tag{10}$$

and

$$B_{H(d)} = \frac{\sum_{S'} \omega_k (h_k - y_k) d_k}{\sum_{S'} \omega_k d_k}. \tag{11}$$

For all adults (i.e., the "domain" where $d_k = 1$ for all adults), each of the three estimators is nearly unbiased. The bias measure for the probability estimator is exactly zero because of how a weighted logistic model is fit, while the bias measure for the standard cut point estimator would also be exactly zero when a perfect cut point satisfying equation (1) can be found. This is also the case for the hybrid cut point estimator when a perfect cut point satisfying equation (7) can be found. Under most models, a perfect cut point cannot be found, but the resulting bias is relatively small.

For all domains (except all adults) because the false positives and false negatives are equal only in expectation, it would be rare for a bias measure to be exactly zero because of random chance. Since each bias measure is a random variable based on the combined MHSS clinical subsample, the question is whether a particular bias measure is *significantly* different from zero.

Each bias measure is a weighted mean across clinical subsample respondents. Its standard error can be estimated (like any weighted mean using the appropriate software[6]) and then put into the denominator of an asymptotically normal t statistic with the bias measure itself as the numerator. This t statistic, for example,

$$t = \frac{B_{P(d)}}{\sqrt{\mathrm{var}(B_{P(d)})}}$$

[5] The bias measure of a model-derived estimator is *not* the difference between it and the direct estimator since a model-derived estimator is computed from the NSDUH sample while its bias is measured in the MHSS subsample.

[6] In our subsequent results, we treat each bias measure as a weighted mean of independent random variables with a common mean, but diverse and unknown individual variances. This ignores the stratification and clustering in the NSDUH main-survey sample. If we had not ignored these aspects of the NSDUH main sampling design, our conclusions would not have changed.

can then be used to test whether the bias measure is significantly different from zero (in this example, the bias measure to be tested is for the probability estimator for domain *d*).

2.3 Bias-Corrected Estimators

The following three *bias-corrected* SMI estimators for domain *d*,

$$\bar{y}_{C(d)-BC} = \bar{y}_{C(d)} - B_{C(d)},$$

$$\bar{y}_{P(d)-BC} = \bar{y}_{P(d)} - B_{P(d)}, \text{ and}$$

$$\bar{y}_{H(d)-BC} = \bar{y}_{H(d)} - B_{H(d)}$$

are, by design, nearly unbiased under probability-sampling theory. In the special case where the domain is all adults, the probability estimator and its bias-corrected analogue are the same.

The standard cut point estimator for all adults only coincides with its bias-corrected analogue when the cut point in the clinical subsample equalizes the weighted number of false positives and false negatives exactly. Otherwise, the two estimators will only be close. Similarly, the hybrid cut point estimator for all adults and its analogue will be the same only when a hybrid cut point satisfying equation (7) can be found.

2.4 Variances and Mean Squared Errors

2.4.1 All Adults

The NSDUH is a stratified multistage sample and the MHSS clinical subsample is a stratified Bernoulli subsample of the NSDUH full sample. As a result, if we ignore the potentially variance-reducing impact of the various calibration weighting steps for the NSDUH full sample, the variance of the probability estimator for all adults under probability-sampling theory is asymptotically the same as the NSDUH main-survey variance of

$$\bar{y} = \frac{\sum_S w_k z_k}{\sum_S w_k}, \tag{12}$$

where

$$z_k = p_k + (\omega_k / w_k)(y_k - p_k).$$ [7]

This makes use of the probability estimator for all adults coinciding with the bias-adjusted probability estimator.

[7] This assumes that **b** in the logistic regression converges to a finite value, say **b***, as the sample size grows arbitrarily large. In the expression for the actual variance of the probability estimator, as opposed to an estimator for that variance, each p_k is replaced by P_k, the value to which it converges.

Similarly, the NSDUH variance of the sample mean in (12) with $z_k = c_k + (\omega_k/w_k)(y_k - c_k)$ estimates the variance of the standard cut point prevalence estimator for all adults in equation (2). When that estimator has a slight bias (i.e., when no perfect cut point can be found), a measure of that bias $\left(\sum_{S'}\omega_k(c_k - y_k)/\sum_{S'}\omega_k\right)$ can be squared and added to the variance to estimate the standard cut point estimator's mean squared error.

The estimated mean squared error of the hybrid cut point estimator for all adults is the estimated variance of the probability estimator plus the square of its bias measure: $B_H = \left(\sum_{S'}\omega_k(h_k - y_k)/\sum_{S'}\omega_k\right)$.

For each of the three model-derived estimates for all adults, an RMSE measure is the square root of its estimated mean squared error.

2.4.2 Domain Level

Because there is either no bias or small bias in the estimator (whether probability or cut point) for SMI prevalence among all adults under probability-sampling theory, it is possible to estimate the prevalence estimator's mean squared error without making assumptions regarding the fit of the model.[8] The same cannot be said about a domain-level estimate. Even when the model holds, the bias measure of a domain-level estimator has only an *expectation* of zero, but it has a positive variance. Consequently, it is not usually zero or nearly zero. As a result, a model-derived, domain-level estimator (whether probability or cut point) cannot be constructed that is nearly equal to an estimator whose variance can be estimated using probability-sampling theory alone.[9] In other words, the simple forms for variance, bias, and RMSE in the previous section do not apply to domain estimators.

Estimating the variance of the model-derived, domain-level estimates using standard linearization methods is difficult because a cut point function (standard or hybrid) is not continuous and cannot be linearized.[10] Only the probability estimator can be linearized.

To circumvent this problem, attempts were made to construct a viable variance measure using a variant of Fay's method of balanced repeated replication (BRR) (Judkins, 1990).[11] Unfortunately, they were not supported by a simulation exercise described below.

---

[8] Again, this assumes only that **b** converges to **b*** as the sample size grows arbitrarily large whether the logistic model holds.

[9] For example, $[\sum_S w_k p_k d_k + \sum_S \omega_k(y_k - p_k)d_k]/\sum_S w_k d_k$ has a variance that can be estimated under probability-sampling theory. Unfortunately, it is *not* nearly equal to the probability estimator $\sum_S w_k p_k d_k / \sum_S w_k d_k$ because $\sum_S \omega_k(y_k - p_k)d_k / \sum_S w_k d_k$ is not exactly or nearly zero even when the model holds.

[10] The probability function $p_k = p(\mathbf{x}_k'\mathbf{b})$ is continuous in $\mathbf{x}_k'\mathbf{b}$, whereas the cut point function changes abruptly from 0 to 1 when $p(\mathbf{x}_k'\mathbf{b})$ first exceeds the cut point. Linearization requires that $p(\mathbf{x}_k'\mathbf{b})$ always have a derivative.

[11] This replication method was chosen because a jackknife is inappropriate for a noncontinuous statistic (see, for example, Judkins, 1990). In addition, BRR is more efficient than the bootstrap when there are two variance primary sampling units (PSUs) per variance stratum as in the MHSS. In fact, bootstrapping can be viewed as random half-sample replication under this design. Balancing maximizes the efficiency of half-sample replication (see, for example, Wolter, 2007). Finally, Fay's version of BRR provides the user with access to the entire sample,

Briefly, because there were two variance primary sampling units (PSUs)[12] within each of 50 variance strata in the MHSS clinical subsample, 52 sets of NSDUH and subsample replicate weights were created using a Hadamard matrix. The matrix dictated which replicate weights were 1.5 times the original weight and which were half of the original weight (with conventional BRR, the replicate weights are either twice the original weight or 0). Each set r (= 1, …, 52) of replicate weights was used to determine an r-specific SMI cut point, and from it an r-specific standard cut point estimate for SMI in a domain. The BRR variance estimate for $\bar{y}_{C(d)}$ was computed using

$$v_{BRR} = \frac{4}{52}\sum_{r=1}^{52}\left[\frac{\sum_S w_{k(r)}c_{k(r)}d_k}{\sum_S w_{k(r)}d_k} - \left(\frac{1}{52}\sum_{\rho=1}^{52}\frac{\sum_S w_{k(\rho)}c_{k(\rho)}d_k}{\sum_S w_{k(\rho)}d_k}\right)\right]^2, \tag{13}$$

where $w_{k(r)}$ was the r replicate weight for NSDUH respondent k, and $c_{k(r)}$ was the respondent's predicted SMI status employing the r-specific SMI cut point.

Using MHSS clinical data from 2008A, 2009, and 2010 with weights computed as described in Center for Behavioral Health Statistics and Quality (CBHSQ, 2014), a simulated NSDUH dataset was created, and 1,679 simulated MHSS clinical subsamples were drawn from it.[13] Standard cut point estimates for SMI among all adults and among Hispanic adults were then computed under the 2008 WHODAS model. Hispanic adults represented a particular domain-level estimate.

Empirical RMSEs were compared with RMSEs computed using the square root of the BRR variance estimator in equation (13). For all adults, the RMSE measure was, on average, roughly 5 percent larger than for empirical RMSE. For Hispanics, it was roughly 23 percent larger.

Because it was likely that the biases of the replicate cut point estimates because of the imperfect cut point were greater than the original, a second variance estimator was computed that removed the biases from the replicate cut point estimates before variance estimation. For all adults, the RMSE measure was, on average, roughly equal to the empirical RMSE, but for adult Hispanics, it remained 19 percent too large. That is, this RMSE measure based on replication worked well for all adults but not for domain estimates.

A model-free RMSE measure for an estimate of SMI prevalence (probability or cut point) among all adults based on linearization was described in Section 2.4.1. The goal in this

---

so that finding a cut point that comes close to equalizing false positives and false negatives in each replicate more likely.

[12] MHSS variance PSUs (an MHVEREP) are usually called "replicates," but we avoid the term here for obvious reasons.

[13] Letting q_k denote the nearest integer to ω_k/w_k, q_k copies of each MHSS respondent were created and randomly assigned to one of the two variance PSUs in the respondent's variance stratum (MHVSTR09). From this simulated NSDUH sample, a simulated Poisson MHSS subsample was drawn using $1/q_k$ as the selection probabilities.

subsection was to see whether the replication variance estimation in equation (13) would lead to a good RMSE measure for domain estimates; it did not.

In Chapter 3, we estimate variances (and from them, RMSE measures are computed) by treating the w_k and ω_k as if they were purely the inverses of selection probabilities and not adjusted for nonresponse or coverage error or to improve estimator efficiency. This may bias the variance estimates upward slightly (more so for the direct estimator than any of the model-derived estimators). The 900 design strata in the NSDUH were collapsed into 50 variance strata with two variance PSUs in each. Because most of the variance of prevalence estimates comes from the subsampling for clinical diagnoses rather than from the stratified NSDUH main sample, the impact from collapsing strata amounts to little more than random noise.[14]

2.5 Evaluating the Models Underlying a Standard Cut Point Estimator

In this section, we describe methods for assessing the logistic model for SMI on which a standard cut point estimator is based. These methods will be used in Chapter 4 in this report to evaluate alternative models we considered for SMI and AMI. The standard cut point estimator is designed so that the (weighted) *false positive rate* in the MHSS clinical subsample, $r_{fp} = \sum_{S';c_k \geq y_k} \omega_k (c_k - y_k) / \sum_{S'} \omega_k$, and the *false negative rate*, $r_{fn} = \sum_{S';c_k < y_k} \omega_k (y_k - c_k) / \sum_{S'} \omega_k$, are as close to equal as possible.

One criterion for comparing the quality of models predicting SMI is the *overall error rate,* $r_e = \sum_{S'} \omega_k \mid c_k - y_k \mid / \sum_{S'} \omega_k$, which measures how well the model correctly predicts an SMI diagnosis. The overall error rate is the sum of the false positive and false negative rates. Models with smaller overall error rates had lower misclassification and therefore were considered superior to models with larger error rates.

The standard cut point estimator is determined in such a way that the false positive and false negative rates for all adults will usually be very close. Unfortunately, that does not mean that the false positive and false negative rates will be close for particular domains of interest such as Hispanic adults or adults having received mental health treatment in the preceding year. In fact, the second bias measure for a standard cut point estimator within a domain in equation (9) is equal to the *difference* between its domain-level false positive and false negative rates. Ideally, a good model for SMI would not produce domain estimates with statistically significant bias measures for domains of interest.

A third criterion for assessing SMI models is the following proxy for the standard cut point estimator's standard error: $\left\{ \sum_{S'} [\omega_k (y_k - c_k)]^2 \right\}^{1/2} \Big/ \sum_{S'} \omega_k$. This measure, which like the overall error rate is computed using only clinical subsample respondents, approximates the square root of the dominant term in the variance estimator for the cut point estimator discussed in

[14] The original intent was to compute standard errors for annual prevalence estimates. To avoid having an empty PSU in 2009, 18 contiguous strata were collapsed into a single variance stratum, making 50 variance strata in all. In retrospect, it may have been better to have used 100 variance strata when analyzing the 2008-2012 MHSS data. Fortunately, preliminary indications are that differences between the standard errors computed using the two sets of variance strata are indeed random.

Section 2.4.[15] It captures only the contribution to a cut point estimator's RMSE because of the clinical subsample being a subsample of the NSDUH main survey. It ignores the bias of the cut point estimator (as a standard error should) and the contribution to variance resulting from the NSDUH main survey also being a probability sample.

2.6 Simultaneously Estimating SMI and AMI with a Cut Point Estimator

One of the consequences of fitting separate logistic models for SMI and AMI and then producing standard cut point estimators for each independently is that the results would likely be logically inconsistent (i.e., some adults would be predicted to have SMI but not AMI). This is possible even if the logistic models for SMI and AMI contain the same predictor variables, since their parameter estimates would almost certainly differ.

Three estimation methods for eliminating this potential inconsistency were evaluated (the numerical results are in Section 6).[16] the *SMI method*, the *AMI method*, and the *combined method*. The first method had been used originally with the 2008 WHODAS model. With this method, SMI is first fit with a weighted logistic model. Then, the WHODAS MHSS clinical subsample S' is sorted by the estimated probabilities of having SMI, $p_k = [1 + \exp(-\mathbf{x}_k'\mathbf{b})]^{-1}$, where $\mathbf{x}_k$ is a vector of predictors, including an intercept. A cut point value p_C is found such that

$$\sum_{\substack{k \in S' \\ p_k \geq p_C}} \omega_k = \sum_{k \in S'} \omega_k y_k \tag{14}$$

holds as nearly as possible. An indicator random variable c_k is then defined for any k in the NSDUH respondent sample to be 1 when $p_k \geq p_C$ and 0 otherwise.

This is the method for determining and applying the SMI cut point discussed in Section 2.1. Under this method, the AMI cut point is determined by using the same sorted probabilities of having SMI but with y_k in equation (1) replaced by a_k, which is defined to be 1 when WHODAS respondent k has AMI and 0 otherwise.

Second, a simple alternative to the *SMI method* for determining logically consistent SMI and AMI cut point estimates is the *AMI method*. It also uses equation (1) for SMI and replaces y_k with a_k for AMI. It differs in that p_k in equation (1) is the estimated probability of k having AMI rather than SMI.

Both the SMI and AMI methods can also be used to determine a cut point for serious or moderate mental illness (SMMI) by defining m_k to be 1 when WHODAS respondent k had

---

[15] The full variance estimator is $\mathrm{var}\left(\sum_S w_k c_k + \omega_k (y_k - c_k)\right) / \sum_S w_k$, where var(.) denotes a variance estimator under probability-sampling theory. The contributions from the variance of $\sum_S w_k c_k / \sum_S w_k$ and the covariance between $\sum_S w_k c_k / \sum_S w_k$ and $\sum_S \omega_k (y_k - c_k) / \sum_S \omega_k$ have been dropped, as have the possible impacts of stratification and clustering in the NSDUH main-survey sample. The weights were constructed so that $\sum_S w_k = \sum_S \omega_k \equiv \sum_{S'} \omega_k$.

[16] An earlier evaluation of alternative methods can be found in Aldworth and colleagues (2010a, pp. 36-38).

SMMI and 0 otherwise and using it in place of y_k in equation (1). SMMI is used in the third method for ensuring logical consistency.

Third, in the *combined method*, the regression coefficients in the models for SMI, AMI, and SMMI are assumed to be the same except for their intercepts.[17] The method creates three data points for each subsample respondent: one with SMI status as the dependent variable, one with SMMI status, and one with AMI status. By using survey sampling software such as SUDAAN®, it is possible to account for these three data points being highly correlated (since they are in the same PSU).

Operationally, the method again employs equation (1) but redefines $\mathbf{x}_k$ and y_k and replaces S′ with a new dataset S″. In creating S″, each respondent along with its MHSS clinical weight generates three data points (so that S″ is three times as large as S′). In the first data point, $y_k = 1$ when k has been diagnosed to have AMI and is 0 otherwise. In the second data point, $y_k = 1$ when k has been diagnosed to have SMMI and is 0 otherwise. In the third data point, $y_k = 1$ when k has been diagnosed to have SMI and is 0 otherwise. The $\mathbf{x}_k$ vector is augmented by two new components that are 0 everywhere except the first component of every first copy and the second component of every second copy is 1. This means that the first new component is a dummy estimating the difference in the log odds of being AMI rather than SMI when all other things (e.g., the remainder of $\mathbf{x}_k$) are equal, while the second new component is a dummy estimating the difference in the log odds of being SMMI rather than SMI when all other thing are equal.[18]

Table 2.1 shows, for example, how the estimated coefficients of the 2012 (WHODAS) model change depending on whether SMI prevalence, AMI prevalence, or joint SMI/AMI/SMMI prevalence are estimated. The weights used in the analysis are described in Section 2.7.

Table 2.1 Logistic Model Coefficients with Different Dependent Variables

	Beta		
Predictor Variables	**Estimate**	**Standard Error**	***p*-value**
SMI as the Dependent Variable			
Intercept	-5.97	0.32	0.0000
Kc_7	0.09	0.02	0.0009
Ac_0	0.34	0.03	0.0000
MHSUITHK	1.96	0.22	0.0000
AMDEYR2	1.13	0.22	0.0000
AGE1830	0.11	0.02	0.0001
AMI as the Dependent Variable			
Intercept	-2.59	0.23	0.0000
Kc_7	0.13	0.02	0.0000
Ac_0	0.33	0.03	0.0000

(continued)

[17] As in a proportional odds model (also called the "ordered logit model"). See McCullagh (1980) and footnote 21.

[18] When the model is correct, fitting a weighted proportional odds model would produce asymptotically identical results, but since pseudo-maximum-likelihood methods were not used, the results would not necessarily be exactly identical. Instead, the method used here forces the predicted estimates of SMI, AMI, and SMMI computed in the clinical subsample to equal the actual subsample estimates of the three.

Table 2.1 Logistic Model Coefficients with Different Dependent Variables (continued)

Predictor Variables	Beta		
	Estimate	Standard Error	*p*-value
MHSUITHK	0.85	0.29	0.0055
AMDEYR2	0.82	0.19	0.0001
AGE1830	0.03	0.02	0.2094
Joint SMI/AMI/SMMI Dependent Variable			
Intercept	-5.14	0.23	0.0000
Kc_7	0.11	0.02	0.0000
Ac_0	0.33	0.02	0.0000
MHSUITHK	1.37	0.23	0.0000
AMDEYR2	0.96	0.16	0.0000
AGE1830	0.03	0.02	0.0955
Dummy for AMI	2.47	0.13	0.0000
Dummy for SMMI	1.31	0.12	0.0000

SMI =serious mental illness; AMI = any mental illness; SMMI = serious or moderate mental illness; Kc_7 = a measure of psychological distress that will be defined fully in Section 4.1; Ac_0 = a measure of functional impairment that will be defined fully in Section 4.1; MHSUITHK = 1 if respondent had serious thoughts of suicide in the past year, 0 otherwise (more in Section 4.1); AMDEYR2 = 1 if respondent had a major depressive episode in the past year, 0 otherwise (more in Section 4.1); AGE1830 = respondent's age minus 18 or 12, whichever is smaller.

NOTE: The *p*-value refers to a two-sided test of whether the estimated regression coefficient (beta) is different from zero.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; Data from 2008A-2012 MHSS clinical respondents. The scaled weights used in combining data across years are described in Section 2.7. Chapter 1 described the 2008A sample.

2.7 Weighting the Clinical Subsample Data

The weights used for the clinical subsample in the 2008 modeling process were developed using simple methods because the sample size was relatively small. The weights were the product of four factors: the respondent's NSDUH weight (for the WHODAS sample), the inverse of the probability that the respondent was selected for the clinical subsample, a nonresponse adjustment, and a "poststratification" adjustment that forced the population estimates computed from the clinical subsample for some demographic groups to match population totals derived from Census Bureau data.

The weighting scheme was revised after extensive analysis of data from 2008 through 2011 (CBHSQ, 2014). This new weighting scheme was used for the entire 2008 to 2012 clinical dataset when evaluating alternative SMI estimation methods and in the final model determination.

The revised single-year weights used in this analysis were the product of six factors: (1) the respondent's NSDUH main-sample weight; (2) a coverage adjustment to account for NSDUH main-study respondents who completed the survey in Spanish; (3) the inverse of the probability the respondent was selected for the clinical subsample; (4) a refusal adjustment to account for NSDUH respondents selected for clinical evaluation who did not wish to be recontacted; (5) a second nonresponse adjustment to account mostly for those who agreed to be recontacted but were unavailable for the clinical evaluation (this included a few who agreed to be recontacted for the evaluations, but refused to respond when recontacted); and (6) a poststratification adjustment.

The revised SMI model was fit using WHODAS subsample data from 2008 through 2012 under the assumption that the model was unchanged across the years (the results of an *F* test was consistent with this assumption). Because the sample size, sampling allocation, and weight adjustments for the subsample differed across the years, gains in statistical efficiency were realized by scaling the weights so that each year of data was not counted equally. The variance of a single estimate that combines data from samples drawn independently across multiple years is minimized when the relative scaling factor attached to each year is proportional to the inverse of the variance of the year's individual estimate.

Because the clinical subsample weights were to be used to compute several direct estimates as well as for modeling, there was not a single set of ideal scaling factors. The scaling factors adopted for modeling were based on a compromise among the optimal estimated variances of annual estimates for SMI, AMI, and past year major depressive episode (MDE). They were 0.12 for 2008, 0.04 for 2009, 0.14 for 2010, 0.35 for 2011, and 0.35 for 2012.[19]

Unless stated otherwise (as in Section 4.1 and 4.2), these were the scaling factors used in the analyses contained in the first six chapters. The scaling factors were applied to *both* the clinical and NSDUH main-survey respondent samples so that the theory underpinning the near unbiasedness of the SMI prevalence estimators would apply. Even then, one needs to either (1) assume the underlying SMI prevalence in a domain stayed constant across the years from 2008 to 2012[20] or (2) treat the target of estimation as the weighted mean of the annual SMI prevalence, where the weight applied to each year is its scaling factor times its relative population size.[21]

Mathematically, the average prevalence from 2008 to 2012 corresponding to assumption (1) can be expressed as

$$\bar{Y} = \frac{N_{2008}\bar{Y}_{2008} + N_{2009}\bar{Y}_{2009} + N_{2010}\bar{Y}_{2010} + N_{2011}\bar{Y}_{2011} + N_{2012}\bar{Y}_{2012}}{N_{2008} + N_{2009} + N_{2010} + N_{2011} + N_{2012}},$$

where N_t and $\bar{Y}_t$ are the adult population size and the SMI prevalence in year t, respectively. The assumption-free target of the scaled estimates (2) is slightly different:

$$\bar{Y}_{scaled} = \frac{(.12)N_{2008}\bar{Y}_{2008} + (.04)N_{2009}\bar{Y}_{2009} + (.14)N_{2010}\bar{Y}_{2010} + (.35)N_{2011}\bar{Y}_{2011} + (.35)N_{2012}\bar{Y}_{2012}}{(.12)N_{2008} + (.04)N_{2009} + (.14)N_{2010} + (.35)N_{2011} + (.35)N_{2012}}.$$

Both these equations apply at the domain level by redefining N_t as the population size of the domain in year t and $\bar{Y}_t$ as the SMI prevalence in that domain and year.

[19] For direct estimates from the clinical subsample using both the 2008A and 2008B samples, the weights for adults in each were scaled by 0.06 (= 0.12/2).

[20] This assumption is *not* contradicted by statistical tests.

[21] Scaling factors are not applied to estimates derived from the NSDUH main-survey data when combining years such as the estimates of SMI based on the SMI model predictions.

3. Choosing an Estimator

In this chapter, we compare the three model-derived estimators for serious mental illness (SMI) discussed in Chapter 2. These are the standard cut point estimator, the probability estimator, and the hybrid cut point estimator. The data used for evaluating the estimators presented here are the WHODAS data from 2008A to 2012. The model used is the 2012 WHODAS model for SMI, which is fully described in Chapter 5. Earlier datasets (e.g., WHODAS data excluding 2012) and models (e.g., the 2008 model) were used in the actual decision-making process before the 2012 model was available. The results were usually very similar.

The probability estimator, which assigns a probability of having SMI to each adult National Survey on Drug Use and Health (NSDUH) respondent, is more efficient (has the smallest root mean square error [RMSE]; see Section 2.3) when estimating SMI prevalence among all adults than the two cut point estimators, which assign a 0/1 SMI status to each adult NSDUH respondent.

The RMSE results are displayed in Table 3.1. All three model-derived estimators had smaller RMSE measures than the direct estimator (which uses only clinical subsample respondents) with the probability estimator having the smallest of all.

Table 3.1 The Estimators for SMI among All Adults Using the 2012 (WHODAS) Model

Estimator	Estimated SMI Prevalence (%)	Bias Measure (%)[1]	RMSE Measure (%)[2]
Direct	3.93	0.00	0.29
Probability	3.91	0.00	0.23
Standard Cut Point	3.95	-0.01	0.26
Hybrid Cut Point	3.89	-0.02	0.26

SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

[1] Bias measure is the weighted mean value of the difference between the true value of SMI and predicted value of SMI taken across the clinical respondent subsample within a particular subpopulation.

[2] RMSE is a measure of the root mean square error. For the standard cut point estimator, it is computed with

$$\left\{\mathrm{var}\left(\sum_S w_k z_k / \sum_S w_k\right) + \left[\sum \omega_k (y_k - c_k) / \sum \omega_k\right]^2\right\}^{1/2},$$

where $z_k = c_k + (\omega_k / w_k)(y_k - c_k)$, and var(.) is a probability-sampling variance estimator computed with MHVSTR09 as the variance strata and MHVEREP as the variance PSUs. For the probability estimator c_k is replaced by p_k. For the hybrid cut point estimator, the c_k within the definition of z_k are replaced by h_k, while the other c_k is replaced by p_k.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 NSDUH adult and MHSS clinical respondents. The scaled weights used in combining data across years were described in Section 2.7. Chapter 1 described the 2008A sample.

As Table 3.2 shows, however, the probability estimator under this model often produced severely biased estimates at the domain level. Under the assumption that an estimator is unbiased, the bias measure of the probability estimator should not be significantly different from zero. The bias measure divided by the standard error of the bias measure is a test statistic for this

assumption. We treat that test statistic here as asymptotically normal under the null hypothesis of no domain-level bias.

Table 3.2 The Standard Cut Point and Probability Estimators for SMI and Their Bias Measures Using the 2012 Model by Domain

Domain	Standard Cut Point Estimator (%)			Probability Estimator (%)		
	Estimate	Bias Measure1	SE of Bias Measure	Estimate	Bias Measure1	SE of Bias Measure
All Adults	3.95	-0.01	0.27	3.91	0	0.23
Male	2.99	0.08	0.42	3.18	0.17	0.34
Female	4.84	-0.10	0.34	4.58	-0.16	0.31
Age: 18-25	3.94	-0.02	0.55	3.59	-0.07	0.49
Age: 26-34	5.03	0.69	0.66	4.64	0.26	0.51
Age: 35-49	5.08	-1.10	0.57[a]	4.77	-1.15	0.55[b]
Age: 50 or Older	2.84	0.37	0.42	3.21	0.61	0.32[a]
White, Not Hispanic	4.31	-0.17	0.33	4.18	-0.16	0.28
Black, Not Hispanic	3.14	-0.48	0.45	3.38	0.00	0.46
Other, Not Hispanic	3.14	-1.14	1.13	3.47	-0.86	1.08
Hispanic	3.31	1.63	0.85[a]	3.28	1.17	0.65[a]
Northeast	3.55	-0.04	0.39	3.62	0.33	0.35
North Central	4.16	0.16	0.60	4.02	-0.10	0.40
South	3.80	-0.13	0.52	3.86	0.22	0.44
West	4.28	0.02	0.56	4.10	-0.56	0.55
Employed Full Time	2.76	0.38	0.33	3.09	0.75	0.28[c]
Employed Part Time	4.19	0.39	0.59	4.05	0.15	0.47
Unemployed	6.61	0.03	0.75	5.48	-0.57	0.70
Other Employment Status	5.33	-0.93	0.66	4.91	-1.30	0.56
Less than High School	4.34	-0.11	0.90	4.15	-0.64	0.83
High School Graduate	4.09	0.01	0.59	3.92	-0.22	0.46
Some College	4.50	0.18	0.37	4.35	0.17	0.31
College Graduate	3.09	-0.16	0.46	3.36	0.33	0.40
Large Metro	3.63	-0.34	0.38	3.68	-0.06	0.35
Small Metro	4.35	0.73	0.49	4.20	0.23	0.35
Nonmetro	4.24	-0.38	0.59	4.09	-0.27	0.51
Health Insurance: Yes	3.67	-0.16	0.27	3.72	0.07	0.24
Health Insurance: No	5.39	0.72	0.86	4.89	-0.34	0.68
< 100% of Poverty Threshold	7.21	-2.07	1.27	6.13	-2.88	1.16[b]
100%-199% of Poverty	4.83	0.12	0.62	4.53	-0.38	0.55
> 200% of the Poverty	2.98	0.32	0.28	3.24	0.61	0.21[c]
Received Mental Health Treatment: Yes	18.33	-1.37	1.31	13.97	-5.07	1.26[c]

(continued)

Table 3.2 The Standard Cut Point and Probability Estimators for SMI and Their Bias Measures Using the 2012 Model by Domain (continued)

Domain	Standard Cut Point Estimator (%)			Probability Estimator (%)		
	Estimate	Bias Measure1	SE of Bias Measure	Estimate	Bias Measure1	SE of Bias Measure
Received Mental Health Treatment: No	1.62	0.20	0.23	2.28	0.81	0.17[c]

SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule; SE = standard error; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study. The 2012 (WHODAS) model will be parameterized in Table 6.1.

[1] Bias measure is the weighted mean value of the difference between the true value of SMI and predicted value of SMI taken across the clinical respondent subsample within a particular subpopulation.

[a] Bias measure is significantly different from zero at the .1 level.

[b] Bias measure is significantly different from zero at the .05 level.

[c] Bias measure is significantly different from zero at the .01 level.

NOTES: *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.
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.
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.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 NSDUH (for the estimates only) and MHSS clinical interview respondents. The scaled weights used in combining data across years are described in Section 2.7. Chapter 1 described the 2008A sample.

One reason the 32 domains Table 3.2 were selected for domain-level bias analyses, outside of the Substance Abuse and Mental Health Services Administration's (SAMHSA's) interest in their prevalence estimates *per se*, was because their relatively large sample sizes and relative homogeneity, making them better candidates for the assessment of bias (test of bias would be more likely to be adequately powered). State-level estimates, which are also of interest to SAMHSA, could not be fruitfully assessed for domain level bias given the sample size of the MHSS clinical dataset.

As previously discussed in Chapter 2, a standard cut point estimator is derived using a cut point that comes as close as possible to equalizing the false positives and false negatives in the clinical subsample. The standard cut point estimator under the 2012 model was significantly biased at the .1 level in only two domains displayed in Table 3.2 and never at the .05 level. Two-sided *p*-values were computed for the 32 domains in the table. Consequently, finding two domains with *p*-values below.1 is consistent with the null hypothesis that the cut point estimator is not biased at the domain level. With 32 significance tests at the .1 level, we would expect to find 3.2 "significant" differences (32 × .1) under the null hypothesis. The actual number of significant differences is only 2.

Table 3.3 repeats much of Table 3.2 for the standard cut point estimator but also displays analogous results for the hybrid cut point estimator. As described in Chapter 2, the cut point was chosen for the hybrid estimator so that the all-adult estimate of SMI prevalence would come as close as possible to matching the probability estimate because the probability estimator for all

adults has a smaller mean squared error than the standard cut point estimator. The hope was that the hybrid cut point estimator would combine the strengths of its two rivals (low mean squared error for all adults with a small amount of domain-level bias).

Table 3.3 The Two Cut Point Estimators for SMI (Standard and Hybrid) and Their Bias Measures Using the 2012 Model by Domain

Domain	Standard Cut Point Estimator (%)			Hybrid Cut Point Estimator (%)		
	Estimate	Bias Measure[1]	SE of Bias Measure	Estimate	Bias Measure[1]	SE of Bias Measure
All Adults	3.95	-0.01	0.27	3.89	-0.10	0.27
Male	2.99	0.08	0.42	2.94	0.03	0.42
Female	4.84	-0.10	0.34	4.78	-0.21	0.33
Age: 18-25	3.94	-0.02	0.55	3.89	-0.03	0.55
Age: 26-34	5.03	0.69	0.66	4.97	0.68	0.66
Age: 35-49	5.08	-1.10	0.57[a]	5.02	-1.16	0.57[b]
Age: 50 or Older	2.84	0.37	0.42	2.79	0.22	0.41
White, Not Hispanic	4.31	-0.17	0.33	4.24	-0.22	0.33
Black, Not Hispanic	3.14	-0.48	0.45	3.10	-0.48	0.45
Other, Not Hispanic	3.14	-1.14	1.13	3.11	-1.14	1.13
Hispanic	3.31	1.63	0.85[a]	3.25	1.30	0.79
Northeast	3.55	-0.04	0.39	3.50	-0.05	0.39
North Central	4.16	0.16	0.60	4.12	0.07	0.59
South	3.80	-0.13	0.52	3.74	-0.29	0.51
West	4.28	0.02	0.56	4.23	0.01	0.56
Employed Full Time	2.76	0.38	0.33	2.71	0.36	0.33
Employed Part Time	4.19	0.39	0.59	4.16	0.37	0.59
Unemployed	6.61	0.03	0.75	6.43	-0.27	0.69
Other Employment Status	5.33	-0.93	0.66	5.27	-1.09	0.65[a]
Less than High School	4.34	-0.11	0.90	4.21	-0.14	0.90
High School Graduate	4.09	0.01	0.59	4.03	-0.26	0.56
Some College	4.50	0.18	0.37	4.45	0.18	0.37
College Graduate	3.09	-0.16	0.46	3.07	-0.17	0.46
Large Metro	3.63	-0.34	0.38	3.58	-0.36	0.38
Small Metro	4.35	0.73	0.49	4.27	0.58	0.47
Nonmetro	4.24	-0.38	0.59	4.19	-0.53	0.58
Health Insurance: Yes	3.67	-0.16	0.27	3.62	-0.20	0.27
Health Insurance: No	5.39	0.72	0.86	5.31	0.44	0.82
< 100% of Poverty Threshold	7.21	-2.07	1.27	7.12	-2.44	1.21[b]
100%-199% of Poverty	4.83	0.12	0.62	4.78	-0.01	0.61
> 200% of the Poverty	2.98	0.32	0.28	2.93	0.30	0.28
Received Mental Health Treatment: Yes	18.33	-1.37	1.31	18.19	-1.46	1.31

(continued)

Table 3.3 The Two Cut Point Estimators for SMI (Standard and Hybrid) and Their Bias Measures Using the 2012 Model by Domain (continued)

Domain	Standard Cut Point Estimator (%)			Hybrid Cut Point Estimator (%)		
	Estimate	Bias Measure[1]	SE of Bias Measure	Estimate	Bias Measure[1]	SE of Bias Measure
Received Mental Health Treatment: No	1.62	0.20	0.23	1.58	0.12	0.22

SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule; SE = standard error; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study. The 2012 (WHODAS) model will be parameterized in Table 6.1.

[1] Bias measure is the weighted mean value of the difference between the true value of SMI and predicted value of SMI taken across the clinical respondent subsample within a particular subpopulation.

[a] Bias measure is significantly different from zero at the .1 level.

[b] Bias measure is significantly different from zero at the .05 level.

[c] Bias measure is significantly different from zero at the .01 level.

NOTE: See Table 3.2 for notes on domain definitions.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 NSDUH (for the estimates only) and MHSS clinical respondents. The scaled weights used in combining data across years were described in Section 2.7. Chapter 1 described the 2008A sample.

The two sets of cut point outcomes were similar, but the bias measure for the hybrid estimator was significantly different from zero at the .05 level in two domains (both with *p*-values of.043). Because 32 domains were analyzed, finding only 2 domains with significant biases at the .05 level remains consistent with the null hypothesis of no bias at the domain level. With 32 significance tests at the .05 level, we would expect to find 1.6 significant differences ($32 \times .05$) under the null hypothesis. The actual number of significant differences is 2, which cannot be viewed as evidence against the null hypothesis.

Even though the null hypothesis of no domain-level bias was consistent with the bias analysis of the hybrid cut point estimator in Table 3.3, there remained a greater *possibility* that the hybrid estimator was biased at the domain level than the standard cut point estimator (2 being greater than 0). Given that there were not many differences between the two estimators and that the standard cut point estimator is easier for data users to calculate, the standard cut point was chosen for use operationally.

The domain-level biases resulting when computing the probability estimator suggest that the logistic model (i.e., the 2012 model on which the estimator is based) may not fit the data. However, computing the standard cut point estimator, which is based on the same model, does not result in the same domain-level biases. This may be because of the ability of the 2012 logistic-model fit to order the probabilities of having SMI correctly. The standard cut point estimators rely only on the order of the probabilities based on the model and not the exact values of these probabilities.

In order to determine whether the probability estimator could be adjusted to account for domain-level biases, attempts were made to use the same ordering as produced by the 2012 model but then to adjust the probabilities assigned to the ordered values in such a way that apparent domain-level biases would disappear. Unfortunately, none of those attempts were

successful in removing significant domain-level biases for SMI among adults receiving mental health treatment or among adults fully employed. One factor complicating the successful assignment of probabilities of having SMI is that over 54 percent of the adult population had the same estimated probability of having SMI: they were over 29 years old, without major depressive episode (MDE) or suicidal thoughts in the previous year, and had low Kessler 6 (K6) and WHODAS scores.

4. Evaluating Alternative Models

As described in Chapter 3, the standard cut point estimator was chosen as the estimator for producing serious mental illness (SMI) and any mental illness (AMI) estimates. An evaluation was simultaneously conducted to determine whether an alternative to the 2008 (World Health Organization Disability Assessment Schedule [WHODAS]) model for SMI could be used to provide more accurate estimates of mental illness, including not only SMI, but AMI and to a lesser extent serious or moderate mental illness (SMMI).

Evaluation of the prediction model began while the 2012 clinical data were still being collected. As a result, various sections of this chapter use different sets of data depending on what was available at the time of the investigation. In Sections 4.1 and 4.2, the prediction-model evaluation used clinical WHODAS data from 2008 through 2011.[22] The goal of the research reviewed in these sections was to select model covariates that produced accurate SMI and AMI cut point estimates for all adults. Specifically, different functions of K6 and WHODAS were investigated, and the inclusion of various mental health indicators in the prediction model were explored.

Section 4.3 describes the analysis conducted with 2008 to 2012 WHODAS data to determine whether bias existed in the mental illness estimates for various domains. The section also describes the minimization of bias observed for age group domains when prediction models employed include an age indicator.

4.1 The Search for New Models

This section describes the variables and recodes of these variables considered for incorporation into a logistic model for SMI (and also for AMI). It then summarizes the steps used in model selection.

4.1.1 Predictor Variables

The 2008 WHODAS model is a logistic model containing two predictor variables: one derived from the National Survey on Drug Use and Health (NSDUH) respondent's Kessler 6 (K6) score and the other from his/her WHODAS score.

The K6 score from the NSDUH main interview, denoted by K6SCMAX, ranges from 0 to 24. The score is based on six measures of psychological distress felt on a respondent's worst 30-day period in the past year ("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, as were "don't knows," refusals, bad data, blanks, and legitimate skips).

[22] For these analyses, the revised yearly Mental Health Surveillance Study (MHSS) clinical weights described in Section 2.7 were scaled for increased statistical efficiency using these factors: 15 percent for MHSS clinical respondents in 2008A, 5 percent for respondents in 2009, 20 percent for respondents in 2010, and 60 percent for respondents in 2011.

The WHODAS functional impairment score from the NSDUH main interview, denoted WHODASC2, also ranges from 0 to 24. It is based on the respondent's level of difficulty performing eight daily activities because of problems with emotions, nerves, or mental health ("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).

We evaluated a rich set of transformations of the K6 and WHODAS scores for possible inclusion within a model for SMI. This involved re-expressing the versions of the K6 and WHODAS measures featured in the 2008 WHODAS model as

- **Kc_7** = max(K6SCMAX – 7, 0)
- **WHODASC3,** which scores each WHODAS task that is either moderately or severely difficult for the respondent to perform with a 1 (as opposed to a 0).[23]

These variables were special cases of transformations in one of the following forms used in our evaluations:

- **Kc_*n*** = max(K6SCMAX – *n*, 0) (*n*= 0, …, 23),
- **Wc_*n*** = max(WHODASC2 – *n*, 0) (*n*= 0, …, 7),
- **Ac_*n*** = max(WHODASC3 – *n*, 0) (*n*= 0, …, 7), and
- **Sc_*n*** = max(WHODASC4 – *n*, 0) (*n*= 0, …, 7),

where WHODASC4 scores only those WHODAS tasks that the respondent had severe difficulty performing with a 1 (and with a 0 otherwise).

The following variables, functions of the preceding four, were also created for convenience:

- **K_*n*** = 1 when K6SCMAX ≥ *n*, 0 otherwise (*n*= 0, …, 23),
- **W_*n*** = 1 when WHODASC2 ≥ *n*, 0 otherwise (*n*= 0, …, 7),
- **A_*n*** = 1 when WHODASC3 ≥ *n*, 0 otherwise (*n*= 0, …, 7),
- **S_*n*** = 1 when WHODASC4 ≥ *n*, 0 otherwise (*n*= 0, …, 7),
- **Kc_*nm*** = min(Kc_*n*, *m*–*n*) (*m* > *n*),
- **Wc_*nm*** = min(Wc_*n*, *m*–*n*) (*m* > *n*),
- **Ac_*nm*** = min(Ac_*n*, *m*–*n*) (*m* > *n*), and
- **Sc_*nm*** = min(Sc_*n*, *m*–*n*) (*m* > *n*).

Observe that a model with, say, the variables Kc_6 and Kc_7 is equivalent to one with the variables K_7 and Kc_7.[24] A model with K_7 alone is equivalent to the true coefficient Kc_6

[23] The original WHODAS task scores are 1 for mildly difficult, 2 for moderately difficult, and 3 for severely difficult (0 otherwise).

[24] Kc_6 = max(K6SCMAX – 6, 0) = I(K6SCMAX ≥ 7) + max(K6SCMAX – 7, 0) = K_7 + Kc_7, where I(A) = 1 if A is true and is 0 otherwise; thus Kc_6 + Kc_7 = K_7 + 2Kc_7. A model with Kc_6 and Kc_7 is equivalent to one with the variables K_7 and Kc_7, while the coefficient for Kc_7 in the latter will be half of the size of its counterpart in the former.

and Kc_7 being equal but with opposite signs. As a result, estimating a model with K_7 and Kc_7 allows one to investigate whether the $|t|$ value of Kc_7 warrants the inclusion of both Kc_6 and Kc_7 in the model (this is why variables like K_7 were created). A second reason for replacing Kc_6 and Kc_7 with K_7 is to assure that the probability of having SMI does not decrease as K6 values increase. Variables of the form Kc_*nm* serve analogous purposes.

A variety of additional variables were considered for possible inclusion into a logistic model for SMI. However, important guidelines were imposed on the number and type of predictor variables that could be included into the model. Only a limited number of covariates could reasonably be added to a logistic model based on the clinical subsample data that have, at most, 100 effective degrees of freedom after collapsing strata in the NSDUH main survey to assure that no primary sampling unit is empty (100 variance strata with two variance primary sampling units each). A useful rule of thumb is that the number of predictor variables should not exceed the square root of the number of variance strata (e.g., 10 [i.e.,$\approx\sqrt{100}$]).[25,26]

Although it seems reasonable that all variables thought to be strongly associated with mental illness should be included in the model, some of these variables were not considered because their inclusion could bias the cut point estimator for a domain of interest. For example, even though having received mental health services is correlated with SMI, using mental health service receipt as a covariate in a model for SMI would produce a cut point that over-predicts SMI prevalence in the domain of adults having received mental health services.

The ultimate goal was to construct a parsimonious model that could be used annually to assess changes in SMI within and between demographic groups without the confounding effects of refitting the model. This was particularly important once the collection of the MHSS clinical subsample ceased. This was another reason why variables related to a domain of interest were avoided. Adding a covariate related to a domain of interest would tend to fix the relationship between SMI and that domain.[27] 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.

With these caveats in mind, additional variables were investigated for possible inclusion in the model because of their known strong correlations with mental illness. These variables included having had serious thoughts of suicide in the previous year (MHSUITHK2), having had a major depressive episode (MDE) in the previous year (AMDEYR2), and ever having had an MDE (AMDELT2).

---

[25] This derives from the notion that the number of primary sampling units (PSUs) minus that number of strata must be much larger 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.

[26] SAMHSA analyzed the clinical subsample in its methodological studies using 50 variance strata (MHVSTR09) each having two variance PSUs because it wanted to be able to analyze each year separately. This resulted in 50 effective degrees of freedom. When the clinical subsample is combined across years, however, it becomes possible to use the 100 variance strata (denoted by MHVESTR), which would result in 100 effective degrees of freedom.

[27] The domains of interest are the domains in Table 3.2.

Missing values for predictor variables were treated as 0s for modeling and prediction purposes. Not only is 0 the most likely value for most of the missing predictor variables used in this model, but also imputing missing predictor values in the clinical subsample and the NSDUH main sample assures that every adult NSDUH respondent will have a predicted SMI status using cut point estimation.

4.1.2 Modeling Steps

SMI was modeled using the following steps:

Piecewise linear functions of the K6 score were first introduced into the model alone, starting with Kc_0 and then Kc_1, Kc_2, and so forth, in turn. Recall from Section 4.1.1 that Kc_*n* = max(K6SCMAX – *n*, 0); n = 0,…,23. Variables with estimated coefficients having |*t*| values greater than 1 were kept in the (weighted) logistic model unless the predicted SMI decreased as the K6 score increased (i.e., the sum of coefficients for Kc_*n* variables was less than 0). Here and in the following steps, variables dropped out of the model when their |*t*| values fell below 1.

Beginning with the selected model (i.e., sets of variables used in the model) after Step 1 (Kc_6 and Kc_7), piecewise linear functions of WHODASC2 (Wc_0, …, Wc_23) were added as appropriate. Again, variables were dropped from the model when the predicted SMI decreased as the WHODASC2 score increased.

Beginning with the selected model after Step 2 (Kc_6, Kc_7, and Wc_1), MDE and suicide variables were added as appropriate.

Step 2 was repeated with piecewise linear functions of WHODASC3 (Ac_0, …, Ac_7) used in place of piecewise linear functions of WHODASC2.

Beginning with the selected model after Step 4 (Kc_6, Kc_7, Ac_0, and Ac_5), MDE and suicide variables were added as appropriate.

Step 2 was repeated with piecewise linear functions of WHODASC4 (Sc_0, …, Sc_7) used in place of piecewise linear functions of WHODASC2.

Beginning with the selected model after Step 6 (Kc_6, Kc_7, Sc_0, and Sc_1), MDE and suicide variables were added as appropriate.

Every function of K6 scores can be expressed as a linear combination of Kc_0 through Kc_23 because K6 scores are restricted to positive integers. This means that every possible function of the K6 can appear on the right-hand side of one of our fitted logistic models. The same thing can be said about functions of the WHODASC2, WHODASC3, and WHODASC4. That said, a model in which the probability of having SMI decreased as the K6, WHODASC2, WHODASC3, and WHODASC4 increased was excluded from consideration since theoretically SMI should be an increasing function of those variables.[28]

Including all three of the new mental health variables—past year suicide thoughts, past year MDE, and lifetime MDE (MHSUITHK2, AMDEYR2, AMDELT2, respectively)—always

[28] The reason why we built the model using piece-wise linear components, like Kc_7, rather than indicators, like K_7 (i.e., recall from Section 4.1.1 that K_*n* = 1 if K6SCMAX ≥ n, and is 0 otherwise; *n* = 0,…,23), is that the former requires fewer exclusions to maintain monotonicity.

tended to improve the logistic fit. Because AMDELT2 has a lifetime reference period and is the only predictor variable not defined exclusively by past year thoughts, symptoms, and experiences, models excluding this variable were also fit.

Models for AMI were also developed using the same steps as for SMI. As with SMI, models in which the predicted AMI decreased as the WHODAS score increased were ruled unacceptable (i.e., excluded from consideration).

4.2 Comparing Models

This section uses the measure developed in Section 2.5 to compare alternative versions of the models for SMI (and AMI) given that a standard cut point estimator would be used to estimate prevalences:

1. The error rate under the standard cut point methodology: $\sum \omega_k \, | y_k - c_k | / \sum \omega_k$.
2. A proxy for the all-adult standard error of the standard cut point estimator: $\dfrac{\left\{ \sum [\omega_k (y_k - c_k)]^2 \right\}^{1/2}}{\sum \omega_k}$.

For both, the smaller the measure, the better the estimation strategy.

The error rate estimates how well the combination of the model and the standard cut point methodology predicts the average adult's SMI status. The error rate is independent of sample size. Therefore, it will not absolutely decrease or increase by the number of years used in its computation (e.g., 2008, or 2008 to 2011, or 2008 to 2012).

The second measure estimates the precision of a cut point estimator for a prevalence rate among all adults. A prevalence rate is a mean, and the standard error of an estimated mean (unlike the error rate) decreases as the sample size increases. In this chapter, we used 2008 to 2011 clinical data for the model evaluations. If we add the 2012 clinical data, the standard errors shown in the following tables will decrease as the sample size will increase.

Table 4.1 displays the (proxy) standard error and error rates for the all-adults SMI cut point estimator under a variety of logistic models for SMI. Each model here is defined by the included predictor variables.

Most of the models displayed in Table 4.1 result from the stepwise search for an SMI model described in Section 4.1.2. For example, after Step 1, the model contains the two predictor variables for SMI: Kc_6 and Kc_7.

Also included in Table 4.1 are

1. a model based on the K6 score (Kc_7) alone;
2. the 2008 WHODAS model, which contains functions of K6 and the WHODAS, i.e. Kc_7 and Ac_0; and
3. the 2008 WHODAS model with AMDELT2, AMDEYR2, and MHSUITHK added.

The models producing the estimators with the smallest standard error and the smallest error rate are in bold font and noted in the table.

Table 4.1 Evaluating the Standard Cut Point Estimators for All Adults Based on Alternative SMI Models

Predictor Variables in the Model	Standard Error of the Estimator (%)	Error Rate of the Estimator (%)
Kc_7 (the pre-2004 model)	0.4370	5.1669
Kc_7 Ac_0 (the refitted 2008 WHODAS model)	0.3888	4.1896
Kc_6 Kc_7	0.4370	5.1669
Kc_6 Kc_7 Wc_1	0.3979	4.2626
Kc_6 Kc_7 Wc_1 AMDELT2 AMDEYR2 MHSUITHK	0.4014	**3.6439**[a]
Kc_6 Kc_7 Wc_1 AMDEYR2 MHSUITHK	0.4023	3.6605
Kc_6 Kc_7 Ac_0 Ac_5	0.3892	4.1783
Kc_6 Kc_7 Ac_0 Ac_5 AMDELT2 AMDEYR2 MHSUITHK	0.3555	3.6779
Kc_6 Kc_7 Ac_0 Ac_5 AMDEYR2 MHSUITHK	0.3590	3.7185
Kc_6 Kc_7 Sc_0 Sc_1	0.3992	4.4238
Kc_6 Kc_7 Sc_0 Sc_1 AMDELT2 AMDEYR2 MHSUITHK	0.4182	3.8448
Kc_6 Kc_7 Sc_0 Sc_1 AMDEYR2 MHSUITHK	0.4146	3.7859
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK	**0.3549**[b]	3.6455

SMI = serious mental illness.

[a] Smallest error rate.

[b] Smallest standard error.

NOTES: See Section 4.1 for the definitions of the variables. The standard error is the proxy measure of standard error.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2011 MHSS clinical respondents. The scaled weights used in combining data across years were described in Section 2.7. Chapter 1 described the 2008A sample.

Table 4.2 mimics Table 4.1 replacing SMI with AMI, both in terms of the model and the all-adults standard cut point estimator. The models resulting from the stepwise search are different (because AMI has been modeled rather than SMI), but the bottom line is nearly the same. The estimator based on the 2008 WHODAS model with AMDELT2, AMDEYR2, and MHSUITHK is either the best or nearly the best in terms of both standard error and error rate.

The results in Tables 4.1 and 4.2 used the stepwise logistic modeling described in Section 4.1.2 as a means to an end: producing better standard cut point estimators. It is important to understand that a better model in terms of the logistic model fit did not necessarily result in a better model for producing a cut point estimator, because the cut point estimator is a combination of the logistic model and a mathematical classification of the predicted values from the logistic model.

The results in Tables 4.1 and 4.2 led to a focus on models for SMI and AMI that contained Kc_7, Ac_0, MHSUITHK, AMDEYR2, and AMDELT2 as the predictor variables.

Table 4.2 Evaluating the Standard Cut Point Estimators for All Adults Based on Alternative AMI Models

Variables in the Model	Standard Error of the Estimator (%)	Error Rate of the Estimator (%)
Kc_7	1.0972	16.6505
Kc_7 Ac_0	0.9574	14.8739
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18	1.0643	16.5621
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 W_1 W_2 Wc_2	1.0105	15.8055
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 W_1 W_2 Wc_2 AMDELT2 AMDEYR2 MHSUITHK	0.9813	14.7593
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 W_1 W_2 Wc_2 AMDEYR2 MHSUITHK	0.9820	14.7364
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 Ac_0 Ac_4	0.9569	14.8348
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 Ac_0 Ac_4 AMDELT2 AMDEYR2 MHSUITHK	0.9530	14.5753
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 Ac_0 Ac_4 AMDEYR2 MHSUITHK	0.9565	14.7427
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 S_1 S_4	1.0426	15.8458
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 S_1 S_4 AMDELT2 AMDEYR2 MHSUITHK	0.9933	15.0075
Kc_24 Kc_46 Kc_68 K_9 Kc_9 Kc_1518 Kc_18 S_1 S_4 AMDEYR2 MHSUITHK	0.9864	14.6795
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK	**0.9440**[a]	**14.4689**[b]

AMI = any mental illness; MHSS = Mental Health Surveillance Study.

[a] Smallest standard error.

[b] Smallest error rate.

NOTES: See Section 4.1 for the definitions of the variables. The standard error is the proxy measure of standard error.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2011 MHSS clinical respondents. Chapter 1 described the 2008A sample.

4.3 Removing Bias at the Domain Level

A good standard cut point estimator for SMI (or AMI) among all adults may nonetheless be biased at the domain level. This section investigates the potential for domain-level bias in mental illness prevalence estimates and proposes a remedy for the apparent bias in age group estimates. No other domain chosen for investigation (see the list in Table 3.3) needed such a remedy to remove significant bias.

Table 4.3 shows that a standard cut point estimator for SMI based on a model with Kc_7, Ac_0, MHSUITHK, AMDEYR2, and AMDELT2 produced significantly biased estimates for the age groups 18 to 25 and 35 to 49. Employing a more parsimonious model that excluded AMDELT2 produced very similar results.

Results from Table 4.3 suggested that adding a variable based on age may reduce this bias. Only continuous age variables were considered because a sharp change in SMI (or AMI) prevalence because of an adult aging a single year was not scientifically plausible. Age variables of the form

$$\mathbf{AGE}\boldsymbol{n} = \text{Max}(\text{AGE} - n, 0)$$

were added to models for SMI and AMI containing Kc_7, Ac_0, MHSUITHK, AMDEYR2, and AMDELT2.

Table 4.3 Some Standard Cut Point Estimators for SMI and Their Bias Measures Using a Model with Kc_7, Ac_0, MHSUITHK, and AMDEYR2

	Standard Cut Point Estimate (%)	Bias Measure (%)	Standard Error of Bias Measure (%)	t-value of Bias Measure	p-value of Bias Measure
Includes AMDELT2					
Age: 18-25	5.75	2.43	0.80	3.05	0.002
Age: 26-34	4.66	0.31	0.65	0.47	0.636
Age: 35-49	4.63	-1.51	0.56	-2.71	0.007
Age: 50 or Older	2.49	-0.04	0.40	-0.11	0.911
Excludes AMDELT2					
Age: 18-25	5.79	2.25	0.81	2.77	0.006
Age: 26-34	4.70	0.42	0.65	0.64	0.521
Age: 35-49	4.62	-1.56	0.58	-2.71	0.007
Age: 50 or Older	2.51	0.03	0.40	0.08	0.936

SMI = serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTES: The *p*-value refers to a two-sided test of whether the bias measure is different from zero. See Section 4.1 for the definitions of the variables.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 NSDUH (for the estimates columns only) and MHSS clinical respondents. The scaled weights used in combining data across years were described in Section 2.7. Chapter 1 described the 2008A sample.

A weighted logistic fit of SMI with Kc_7, Ac_0, MHSUITHK, AMDEYR2, AMDELT2, and AGE18 as the predictor values resulted in AGE18 having a |t| value greater than 1, so it was kept in the model. Each AGE*n* variable, $n = 19, 20, \ldots$ was added to the model in turn and assessed. Some of the additions also had |t| values greater than 1. Among them, adding the AGE37 variable to the model (along with Kc_7, Ac_0, MHSUITHK, AMDEYR2, AMDELT2, and AGE18) produced the lowest sum of squared residuals.

Using a parallel procedure in modeling AMI led to adding AGE18 and AGE30 to the model with Kc_7, Ac_0, MHSUITHK, AMDEYR2, and AMDELT2.[29]

In the logistic model for SMI, the coefficients on AGE18 and AGE37 had opposite signs but similar magnitudes: the former positive, the latter negative. The same was true for coefficients on AGE18 and AGE30 in the model for AMI. This suggested we consider a more parsimonious predictor variable of the form:

$$\mathbf{AGE}\boldsymbol{nm} = \mathrm{Min}(\mathrm{AGE} - n,\ m - n).$$

[29] The same additions were made when modeling serious of moderate mental illness (SMMI).

Including this variable, say, AGE1830, in a model for SMI would mean that only adults between 18 and 30 would have their predicted SMIs affected by age, all other predictors kept constant. After 30, predicted SMI would no longer be a function of age.

Table 4.4 shows how the accuracy of an all-adult cut point estimator varied with alternative choices for the age variables in a model that otherwise contains the same predictor variable (Kc_7, Ac_0, MHSUITHK, AMDEYR2, and AMDELT2). The best choice for the SMI is AGE1837. For AMI, it is AGE1830.

Table 4.4 Evaluating SMI and AMI Cut Point Estimators for All Adults Employing Different Variables in the Model

Predictor Variables in the Model	Standard Error of Estimator (%)	Error Rate of Estimator (%)
For SMI		
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK	0.3549	3.6455
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE18	0.3556	3.6475
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE18 AGE37	**0.3237**[b]	**3.3146**[a]
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE1837	0.3318	3.4122
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE18 AGE30	0.3282	3.3389
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE1830	0.3309	3.3730
For AMI		
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK	0.9440	14.4689
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE18	0.9436	14.5160
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE18 AGE37	0.9660	14.3951
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE1837	0.9338	14.0958
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE18 AGE30	0.9665	14.4328
Kc_7 Ac_0 AMDELT2 AMDEYR2 MHSUITHK AGE1830	**0.9334**[b]	**14.0227**[a]

SMI = serious mental illness; AMI = any mental illness; MHSS = Mental Health Surveillance Study.

[a] Lowest error rates.

[b] Lowest standard errors.

NOTES: See Sections 4.1 and 4.3 for the definitions of the variables. The standard error is the proxy measure of standard error.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2011 clinical respondents. The scaled weights used in combining data across years were described in Section 2.7. Chapter 1 described the 2008A sample.

This page intentionally left blank

5. The 2012 Models and Estimators

This chapter describes the final models and estimators chosen for predicting serious mental illness (SMI) and other mental-illness prevalence among adults based on the 2008 to 2012 National Survey on Drug Use and Health (NSDUH). Section 5.1 discusses the estimation methods for respondents who were administered the World Health Organization Disability Assessment Schedule (WHODAS), while Section 5.2 describes how a comparable model was developed for the subsample of respondents who were administered the Sheehan Disability Scale (SDS) in 2008. Analyses in this chapter were based on data from the 2008 to 2012 MHSS clinical subsamples and the 2008 to 2012 NSDUHs.

5.1 The 2012 WHODAS Estimation Method

The final result from previous analyses using the 2008A to 2012 clinical data (described in Chapters 3 and 4) was to use the cut point estimator for models of SMI and AMI, and to include covariates of Kc_7, Ac_0, MHSUITHK, and AMDEYR2 as predictors in the model. It also was determined that AMDELT2 (lifetime major depressive episode [MDE]) would be dropped from further consideration partially for parsimony and partially because all the other predictor variables in the model are related to the respondent's experiences in the previous year rather than his/her lifetime.

To remove the potential for bias in age group prevalence estimates in a relatively parsimonious manner, it was determined that an age variable or combination of age variables would be added to the model. Specifically, a variable termed AGE1830[30] was considered as a potential addition to the model as well as including separate age variables (named AGE18 and AGE30) in the model. Both of these options led to a reduction in bias for SMI and AMI estimates by age.[31]

Unlike some of the analyses in the previous chapter (Section 4.2), the analyses in this section include data from 2012. Also, all models in this chapter include age variables as well as the variables on past year suicidal thoughts and MDE (MHSUITHK and AMDEYR2) since they were previously found to be good predictors of mental illness.

Table 5.1 displays weighted logistic models for both SMI and AMI using the 2008A to 2012 data and combinations of the predictor variables AGE1830, AGE30, Kc_7, Ac_0, MHSUITHK, and AMDEYR2. Including AGE1830 and AGE30 in the model is equivalent to including AGE18 and AGE30 but allows us to test whether AGE18 and AGE30 have equal and opposite coefficients. That is, previous results determined that the probability of SMI or AMI increases up to age 30 but further exploration was needed to determine if the increase simply stops at 30 or whether the probability of having SMI or AMI starts to decrease slightly as an adult ages past 30.

[30] AGE1830 would be a more parsimonious addition to the model than AGE1837, since (unlike AGE1837) no one between 30 and 37 would be affected by its addition.

[31] The results of adding AGE1830 can be seen in Table 3.2.

Table 5.1 Coefficients of Logistic Models for SMI and AMI with AGE1830 and AGE30 as Predictor Variables

Predictor Variables	Beta			
	Estimate (%)	Standard Error (%)	*t*-value	*p*-value
SMI as the Dependent Variable				
Intercept	-6.03	0.31	-19.22	0.0000
Kc_7	0.08	0.03	3.17	0.0026
Ac_0	0.34	0.03	9.76	0.0000
MHSUITHK	1.98	0.21	9.24	0.0000
AMDEYR2	1.17	0.22	5.28	0.0000
AGE1830	0.13	0.03	4.86	0.0000
AGE30	-0.01	0.01	-1.50	0.1408
AMI as the Dependent Variable				
Intercept	-2.66	0.24	-10.93	0.0000
Kc_7	0.12	0.02	6.53	0.0000
Ac_0	0.33	0.03	10.38	0.0000
MHSUITHK	0.88	0.29	3.01	0.0041
AMDEYR2	0.85	0.19	4.35	0.0001
AGE1830	0.05	0.02	2.07	0.0441
AGE30	-0.01	0.01	-1.11	0.2739

SMI = serious mental illness; AMI = any mental illness; MHSS = Mental Health Surveillance Study.

NOTES: The *p*-value refers to a two-sided test of whether the estimated regression coefficient (beta) is different from zero.

See Section 4.1 for the definitions of the variables.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 MHSS clinical respondents. The scaled weights used in combining data across years were described Section 2.7. Chapter 1 described the 2008A sample.

As Table 5.1 shows, the coefficient for AGE1830 is significant at the .05 level for both SMI and AMI. This indicates that when all other predictor variables are held constant, the probability of having either SMI or AMI increases with age from 18 to 30. To test whether the probability of having SMI or AMI only stops increasing at age 30 or whether it starts to decrease slightly as an adult ages past 30, the variable AGE30 was added to the model. Examining the estimated coefficient for the AGE30 variable shows that it is negative but not significant at the .1 level for either SMI or AMI. Nevertheless, the $|t|$ value of the estimated coefficients is greater than 1, indicating that there may be some efficiency gains from including it in the models and a slight potential for bias from excluding it.

The goal, however, is not fitting the best logistic model for SMI or AMI *per se*. It is to estimate SMI and AMI prevalences in a consistent manner with a reasonably well-performing standard cut point estimator. In section 2.6, three ways to create standard cut point estimates were described that ensured that no adult NSDUH main-survey respondent is predicted to have SMI but not AMI, SMI but not SMMI, or SMMI but not AMI. One method is to base the cut points for SMI, AMI, and serious of moderate mental illness (SMMI) on a logistic model for SMI.[32] A second is to base the cut points for SMI, AMI, and SMMI on a logistic model for AMI.

[32] Recall that this means that within the clinical subsample sorted by the estimated probabilities of having SMI, cut points are established that come as close as possible to balancing the weighted number of adults falsely classified as (1) having SMI and not have SMI (the SMI cut point), (2) having and not having AMI (the AMI cut point), and (3) having and not having SMMI (the SMMI cut point).

A third is to base the three cut points on a multivariate model for SMI, AMI, and SMMI that has the same coefficients for each predictor but separate dummies for SMI, AMI, and SMMI. These three methods are referred to as cut point estimation methods in the following text.

Tables 5.2, 5.3, and 5.4 provide output summaries for the three different cut point estimation methods under a variety of models (i.e., sets of predictor variables). This output was used to address the following questions:

1. Were there meaningful efficiency gains (decreases in root mean squared error) from adding MHSUITHK and AMDEYR2 to the model as suggested in the previous chapter, especially the model containing age variables but not MHSUITHK and AMDEYR2?
2. Should the age variables AGE18, AGE30, and/or AGE1830 be added to the model?
3. Which cut point estimation method should be used to assure logical consistency in categorizing a person's SMI/AMI/SMMI status?

These three tables display the standard cut point estimates computed both with scaled weights and with not-scaled weights for the NSDUH adult full sample.[33] The estimates computed with scaled weights almost always come very close to the direct estimates computed from the clinical WHODAS subsample with the scaled weights for the clinical data (3.93 for SMI and 18.03 for AMI). The exceptions occur under the models with only the Kc_7 variable as a predictor since there are only 16 realized values each for the estimated probabilities of having SMI and AMI, which makes it very hard to nearly equalize false positives and false negatives, whatever the cut point method.

The standard cut point estimates computed with not-scaled weights are what would be published for the years 2008 to 2012 taken as a whole (with the SDS sample information from 2008B excluded) if that cut point method and model were used.[34] These unscaled estimates also tend to be close to the direct estimates. This supports the assumption that there is not an underlying difference in the all-adults prevalence of SMI and AMI across the years 2008 to 2012 (otherwise, the direct estimates, which scale the years from 2008 to 2012, and "unscaled" estimates, in which the years are not scaled, could be very different).

The tables show that error rate and root mean square error (RMSE) measure for the SMI prevalence estimate among all adults were smaller when MHSUITHK and AMDEYR2 were added to the model, regardless of the cut point estimation method used. This was mostly the case for the AMI prevalence estimate as well, but the reduction in the error rate and RMSE from adding these two variables was smaller. Since determining a method to optimally estimate SMI

[33] Published cut point estimates are not computed with scaled weights because the annual NSDUH samples were large and roughly constant in size from 2008 through 2012. The error measures in the tables, however, are based on the clinical subsample using scaled weights. As a results, cut point estimates are displayed in the tables using both scaled and not-scaled weights.

[34] Scaling the annual clinical subsamples reduced the variance of estimates produced from them. Scaling the NSDUH main-survey weights analogously render measures of bias and RMSE cleaner in that they are strictly because of the estimation method and not differences in how the weights are scaled. When computing prevalence estimates from the NSDUH main-survey sample for public release, however, NSDUH main-survey weights are not scaled (not scaling the weights is equivalent to scaling them each by the same factor).

is the primary goal of the Mental Health Surveillance Study (MHSS), it was decided to include MHSUITHK and AMDEYR2 within the 2012 WHODAS model (along with Kc_7 and Ac_0 from the 2008 WHODAS model).

To assess the potential for bias under a particular model, Tables 5.2, 5.3, and 5.4 display the average t^2 value among the 32 bias measures (one for each domain in Table 3.2). Ideally, this average should be 1 or smaller if there were no domain-level biases.

Since there are *p*-values for 32 domain-level bias measures, even in the absence of any domain-level bias, it would not be unusual for one to appear statistically significant because of random chance. This is because a significant difference at the .05 level means there is 1 in 20 chance that a result will appear to be significant when it is not. That is why the tables display the *second* lowest *p*-value among the 32 (which is greater than 20). In the absence of any domain-level bias, it would ideally be greater than .05 (because we expect only 1.6 *p*-values to be less than .05).

When the model with SMI as the dependent variable was used to create the cut points (this is the SMI method displayed in Table 5.2) and the predictor variables were Kc_7, Ac_0, MHSUITHK, AMDEYR2, and AGE1830, the estimates for both SMI and AMI had mean t^2- values less than 1 and second lowest *p*-values no smaller than .05, which are consistent with the absence of domain-level bias. When other estimation methods (the AMI or combined methods displayed in Tables 5.3 and 5.4, respectively) were used or when the SMI method was used without an age-related predictor variable, the second lowest *p*-values for either SMI or AMI were less than .05. Thus, it was decided to continue using the SMI cut point method along with the 2012 model.

Table 5.2 reveals few meaningful differences between using the SMI cut point method with a model featuring Kc_7, Ac_0, MHSUITHK, AMDEYR2, AGE18, and AGE30 as the predictor variables and replacing the two age variables with the more parsimonious AGE1830. The former version has a lower error rate and RMSE for SMI, the latter for AMI. Both differences are very small. Ultimately, the relative parsimony of the model with only one age variable added, AGE1830, was the deciding factor in its selection for estimating mental illness prevalence.

Table 5.2 Summary Output of Standard Cut Point Estimates Using the Model Predicting SMI (SMI Method)

Predictor Variables	Cut Point Estimate with Scaled Weights (%)	Cut Point Estimate with Not-scaled Weights (%)	Error Rate (%)	False Positive Rate (%)	False Negative Rate (%)	RMSE (%)	Mean t^2-value for the Bias Measures	2nd Lowest *p*-value for the Bias Measures
For SMI								
Kc_7	4.40	4.28	5.31	2.76	2.55	0.393	2.265	0.000
Kc_7 Ac_0	3.97	3.89	4.42	2.21	2.21	0.289	1.180	0.009
Kc_7 Ac_0 AGE18 AGE30	3.84	3.76	4.30	2.16	2.14	0.281	0.940	0.062
Kc_7 Ac_0 AGE1830	3.89	3.80	4.30	2.17	2.12	0.286	0.900	0.050
Kc_7 Ac_0 MHSUITHK AMDEYR2	3.90	3.83	3.97	1.99	1.98	0.279	1.162	0.007
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	4.01	3.95	3.76	1.88	1.88	0.265	0.854	0.076
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830	3.95	3.89	3.84	1.92	1.93	0.269	0.808	0.056
For AMI								
Kc_7	17.49	17.51	17.43	8.43	8.99	1.313	1.157	0.034
Kc_7 Ac_0	17.93	17.79	15.85	7.99	7.86	1.123	1.116	0.009
Kc_7 Ac_0 AGE18 AGE30	17.98	17.85	15.50	7.75	7.75	1.104	0.851	0.041
Kc_7 Ac_0 AGE1830	17.96	17.83	15.58	7.79	7.79	1.114	0.828	0.051
Kc_7 Ac_0 MHSUITHK AMDEYR2	18.29	18.23	15.77	7.92	7.85	1.109	1.154	0.014
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	18.00	18.01	15.62	7.78	7.84	1.119	0.907	0.046
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830	18.11	18.06	15.47	7.70	7.77	1.116	0.827	0.052

SMI = serious mental illness (the direct estimate for the clinical subsample using scaled weights is 3.93); RMSE = root mean square error; AMI = any mental illness (the direct estimate form the clinical subsample is 18.03). See Chapter 4 for definitions of the model variables.

NOTES: RMSE is a measure of the root mean square error of the standard cut point estimator for all adult computed with scaled weights:

$$\left\{\mathrm{var}\left(\sum_S w_k z_k / \sum_S w_k\right) + \left[\sum \omega_k (y_k - c_k) / \sum \omega_k\right]^2\right\}^{1/2},$$

where $z_k = c_k + (\omega_k / w_k)(y_k - c_k)$, and var(.) is a probability-sampling variance estimator computed with MHV09STR as the variance strata and MHVEREP as the variance PSUs.

Mean t^2-value is the average of the squared *t*-value among the bias measures for 32 domain estimates.

p-value is the two-sided *p*-value under the null hypothesis that the bias measure is estimating zero.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 MHSS clinical respondents. The scaled weights used in combining data across years were described Section 2.7. Chapter 1 described the 2008A sample.

Table 5.3 Summary Output of Standard Cut Point Estimates Using the Model Predicting AMI (AMI Method)

Predictor Variables	Cut Point Estimate with Scaled Weights (%)	Cut Point Estimate with Not-scaled Weights (%)	Error Rate (%)	False Positive Rate (%)	False Negative Rate (%)	RMSE (%)	Mean t^2-value for the Bias Measures	2nd Lowest *p*-value for the Bias Measures
For SMI								
Kc_7	4.40	4.28	5.31	2.76	2.55	0.393	2.265	0.000
Kc_7 Ac_0	3.99	3.88	4.42	2.23	2.19	0.293	1.218	0.007
Kc_7 Ac_0 AGE18 AGE30	3.91	3.83	4.35	2.18	2.18	0.287	1.025	0.035
Kc_7 Ac_0 AGE1830	3.86	3.76	4.37	2.19	2.18	0.288	1.039	0.042
Kc_7 Ac_0 MHSUITHK AMDEYR2	3.98	3.88	4.18	2.09	2.09	0.288	1.330	0.013
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	4.04	3.94	4.16	2.08	2.08	0.285	1.078	0.033
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830	3.99	3.89	4.21	2.10	2.10	0.289	1.052	0.027
For AMI								
Kc_7	17.49	17.51	17.43	8.43	8.99	1.313	1.157	0.034
Kc_7 Ac_0	18.16	18.02	15.94	8.19	7.76	1.194	1.274	0.006
Kc_7 Ac_0 AGE18 AGE30	17.94	17.86	15.74	7.88	7.86	1.114	0.999	0.047
Kc_7 Ac_0 AGE1830	17.88	17.75	15.63	7.84	7.79	1.116	0.918	0.046
Kc_7 Ac_0 MHSUITHK AMDEYR2	17.65	17.57	15.15	7.42	7.74	1.159	1.071	0.018
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	17.85	17.82	15.30	7.65	7.65	1.104	1.015	0.019
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830	18.10	17.99	15.19	7.72	7.47	1.129	1.030	0.017

AMI = any mental illness (the direct estimate for the clinical subsample using scaled weights is 18.03); RMSE = root mean square error; SMI = serious mental illness (the direct estimate for the clinical subsample is 3.93). See Chapter 4 for definitions of the model variables.

NOTES: RMSE is a measure of the root mean square error of the standard cut point estimator for all adult computed with scaled weights:

$$\left\{\mathrm{var}\left(\sum_S w_k z_k / \sum_S w_k\right) + \left[\sum \omega_k (y_k - c_k) / \sum \omega_k\right]^2\right\}^{1/2},$$

where $z_k = c_k + (\omega_k / w_k)(y_k - c_k)$, and var(.) is a probability-sampling variance estimator computed with MHV09STR as the variance strata and MHVEREP as the variance PSUs.

Mean t^2-value is the average of the squared *t*-value among the bias measures for 32 domain estimates.

p-value is the two-sided *p*-value under the null hypothesis that the bias measure is estimating zero.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 MHSS clinical respondents. The scaled weights used in combining data across years are described in the beginning of Section 2.7. Chapter 1 described the 2008A sample.

Table 5.4 Summary Output of Standard Cut Point Estimates Using a Multivariate Prediction Model (Combined Method)

Predictor Variables	Cut Point Estimate with Scaled Weights (%)	Cut Point Estimate with Not-scaled Weights (%)	Error Rate (%)	False Positive Rate (%)	False Negative Rate (%)	RMSE (%)	Mean t^2-value for the Bias Measures	2nd Lowest p-value for the Bias Measures
For SMI								
Kc_7	4.40	4.28	5.31	2.76	2.55	0.393	2.265	0.000
Kc_7 Ac_0	4.05	3.95	4.44	2.26	2.19	0.300	1.273	0.009
Kc_7 Ac_0 AGE18 AGE30	3.94	3.86	4.36	2.18	2.18	0.288	1.048	0.040
Kc_7 Ac_0 AGE1830	3.92	3.82	4.40	2.20	2.20	0.291	1.096	0.026
Kc_7 Ac_0 MHSUITHK AMDEYR2	3.93	3.84	3.97	1.98	1.99	0.275	1.167	0.004
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	3.92	3.85	3.88	1.94	1.94	0.270	0.957	0.036
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830	3.91	3.83	3.96	1.98	1.98	0.271	0.928	0.058
For AMI								
Kc_7	17.49	17.51	17.43	8.43	8.99	1.313	1.157	0.034
Kc_7 Ac_0	17.93	17.79	15.85	7.99	7.86	1.123	1.116	0.009
Kc_7 Ac_0 AGE18 AGE30	17.87	17.80	15.66	7.83	7.83	1.113	1.043	0.036
Kc_7 Ac_0 AGE1830	17.88	17.74	15.64	7.85	7.79	1.117	0.918	0.045
Kc_7 Ac_0 MHSUITHK AMDEYR2	18.46	18.38	15.52	7.89	7.63	1.133	1.236	0.006
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	18.19	18.19	15.36	7.68	7.68	1.102	1.043	0.016
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830	18.10	18.02	15.21	7.60	7.61	1.101	0.948	0.025

SMI = serious mental illness (the direct estimate for the clinical subsample is 3.93); RMSE = root mean square error; AMI = any mental illness (the direct estimate for the clinical subsample is 18.03); NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study. See Chapter 4 for definitions of the model variables.

NOTES: RMSE is a measure of the root mean square error of the standard cut point estimator for all adult computed with scaled weights:

$$\left\{\mathrm{var}\left(\sum_S w_k z_k / \sum_S w_k\right) + \left[\sum \omega_k (y_k - c_k) / \sum \omega_k\right]^2\right\}^{1/2},$$

where $z_k = c_k + (\omega_k / w_k)(y_k - c_k)$, and var(.) is a probability-sampling variance estimator computed with MHV09STR as the variance strata and MHVEREP as the variance PSUs.

Mean t^2-value is the average of the squared t-value among the bias measures for 32 domain estimates.

p-value is the two-sided p-value under the null hypothesis that the bias measure is estimating zero.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 NSDUH and MHSS clinical respondents. The scaled weights used in combining data across years, except when specified otherwise, are described in Section 2.7. Chapter 1 described the 2008A sample.

5.2 The 2012 Model for the 2008 SDS Sample

As described in Chapter 1, roughly half of the adult NSDUH main sample in 2008 was assigned functional-impairment questions based on the SDS rather than the WHODAS. A subsample of these adults was assigned to receive a clinical interview identical to that offered the WHODAS clinical subsample.

The NSDUH main-survey adult sample in 2008 who were assigned SDS-based questions is referred to as the 2008B sample (as opposed to the 2008A sample). The original 2008 SDS model for SMI fit on the clinical subsample contained a recoded SDS total score variable. When that model was applied to the 2008B NSDUH main-survey adult sample to produce an AMI prevalence estimate for all adults, the result (16.73 percent) was lower than and not comparable to the AMI estimates for 2008A and 2009 calculated using the 2008 WHODAS model (19.48 and 19.86 percent[35]). See Aldworth and colleagues (2010b, pp. 36-38). As a result, it was recommended that AMI (and SMMI) prevalence estimates *not* use data from the 2008B sample. See, for example, the 2011 NSDUH mental health findings report (Center for Behavioral Health Statistics and Quality [CBHSQ], 2012).

We investigated whether the recommendation could be reversed with a 2012 model. To that end, the 2008B clinical subsample was weighted using the same methods described in Section 2.7, the entire 2008 to 2012 clinical subsample was fit with a model identical to the 2012 WHODAS model except that the WHODAS variable (Ac_0) was removed. In addition, for the 2008B respondents, an adjusted version of AMDEYR2 was used in its place.[36] Cut points for SMI, AMI, and SMMI were determined from this model and applied to the 2008B NSDUH main survey adult sample respondents only. (Results for SMMI are not shown.)

Tables 5.5 and 5.6 show that there are only small differences between domain and overall estimates of SMI and AMI for the 2008A WHODAS NSDUH respondent sample (estimates based on the 2012 WHODAS model), the 2008 to 2012 clinical subsample (direct estimates), and the 2008B SDS NSDUH respondent sample (estimates based on the 2012 SDS model). The tables contain comparisons between the 2008A (WHODAS) and 2008B (SDS) NSDUH respondent samples, as well as the statistics on the bias measure for the 2008B sample estimates. No bias measure is significantly different from zero at the .1 level.

Tables A.1 and A.2 in Appendix A show what would have happened if SMI and AMI predicted values for the 2008B sample were made from a model for SMI analogous to the 2012 WHODAS model but with an SDS analogous to AC_0 replacing AC_0. The model was then fit to the 2008B sample (the only NSDUH main-survey sample to which SDS-derived questions were assigned). The results are not as consistent with the 2008A results as the 2012 model ultimately chosen for use with the 2008B sample. Consequently, the recommendation to use a model without an SDS term was confirmed.

---

[35] These estimates were made before subsequently discovered falsified records. See Liao and colleagues (2012).

[36] The variable AJAMDEYR was developed to remove the context effects on MDE in the 2008B NSDUH instrument. See Aldworth and colleagues (2012).

Table 5.5 Comparing Standard Cut Point Estimates for SMI from the 2008B Sample Using the 2012 Model for the SDS Sample with 2012 Method Estimates from the 2008A Sample by Domain

Domain	Standard Cut Point Estimates (%)			Bias Measures for 2008B Estimates (%)		
	2008A	2008B	Difference	Measure	Standard Error	*p*-Value
All Adults	3.72	3.71	0.01	-2.54	1.80	0.158
Age: 18-25	4.15	2.62	1.53	-1.15	0.85	0.175
Age: 26-34	4.57	4.88	-0.31	-0.69	2.16	0.750
Age: 35-49	4.74	4.84	-0.10	-0.91	1.62	0.572
Age: 50+	2.52	2.86	-0.34	-4.90	4.14	0.236
White, Not Hispanic	4.14	4.25	-0.11	-2.42	2.47	0.326
Black, Not Hispanic	2.66	3.01	-0.35	-3.01	4.37	0.491
Other, Not Hispanic	3.40	1.93	1.47	-7.82	5.23	0.135
Hispanic	2.58	2.42	0.16	-0.24	0.25	0.337
Male	2.76	2.35	0.41	-1.94	1.29	0.134
Female	4.61	4.98	-0.37	-3.10	3.24	0.339
Northeast	3.36	4.05	-0.69	-5.15	5.49	0.348
Midwest	4.23	3.65	0.58	0.41	0.64	0.525
South	3.71	3.46	0.25	-2.88	2.10	0.170
West	3.53	3.90	-0.37	-1.46	2.24	0.514
Large Metro	3.55	3.53	0.02	0.06	0.77	0.941
Small Metro	3.82	4.09	-0.27	-4.52	4.58	0.324
Nonmetro	4.07	3.61	0.46	-4.46	3.05	0.143
Less Than High School	3.28	3.74	-0.46	-9.75	7.11	0.170
High School Graduate	4.38	3.61	0.77	0.52	0.62	0.409
Some College	4.33	4.73	-0.40	-5.75	5.48	0.294
College Graduates	2.65	2.89	-0.24	-0.95	1.18	0.421
Employed Full Time	2.69	3.00	-0.31	-1.05	1.04	0.315
Employed Part time	4.06	2.97	1.09	-1.56	1.79	0.385
Unemployed	7.98	5.40	2.58	6.05	4.31	0.161
Others	5.00	5.17	-0.17	-5.99	5.07	0.237
Health Insurance: Yes	3.54	3.43	0.11	-1.91	1.89	0.311
Health Insurance: No	4.71	5.31	-0.6	-9.05	5.80	0.119
<100% Poverty Threshold	6.25	6.11	0.14	-17.81	11.33	0.116
100%-199% Poverty Threshold	4.84	4.98	-0.14	-0.37	1.24	0.767
≥200% Poverty Threshold	3.02	3.00	0.02	-0.07	0.64	0.915
Received Mental Health Treatment: Yes	18.39	19.41	-1.02	-16.87	11.32	0.136
Received Mental Health Treatment: No	1.43	1.28	0.15	-0.32	0.47	0.504

SMI = serious mental illness; SDS = Sheehan Disability Scale; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

Bias measure is the weighted mean value of the difference between the true value of SMI and predicted value of SMI taken across the clinical respondent subsample within a particular subpopulation.

NOTE: See Table 3.2 for notes on domain definitions.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. Chapter 1 described the 2008A and 2008B samples.

Table 5.6 Comparing Standard Cut Point Estimates for AMI from 2008B Sample Using the 2012 Model for the SDS Sample with 2012 Method Estimates from the 2008A Sample by Domain

Domain	Standard Cut Point Estimates (%)			Bias Measures for 2008B Estimates (%)		
	2008A	2008B	Difference	Measure	Standard Error	*p*-value
All Adults	17.69	17.46	0.23	-0.14	2.78	0.960
Age: 18-25	18.83	16.36	2.47	-3.87	3.32	0.244
Age: 26-34	20.75	21.68	-0.93	-4.92	5.54	0.375
Age: 35-49	20.08	21.13	-1.05	4.09	5.02	0.415
Age: 50+	14.43	13.66	0.77	0.53	5.31	0.920
White, Not Hispanic	18.44	17.73	0.71	-1.57	3.45	0.649
Black, Not Hispanic	15.79	18.07	-2.28	-5.90	7.20	0.413
Other, Not Hispanic	16.70	16.55	0.15	16.01	9.88	0.105
Hispanic	15.91	15.99	-0.08	4.27	7.40	0.564
Male	13.45	13.93	-0.48	4.27	3.90	0.273
Female	21.64	20.74	0.90	-4.25	3.90	0.276
Northeast	16.68	19.97	-3.29	1.14	7.55	0.880
Midwest	18.64	17.64	1.00	4.93	2.49	0.047
South	17.50	16.16	1.34	-7.07	3.85	0.066
West	17.89	17.31	0.58	3.24	6.88	0.637
Large Metro	16.91	17.31	-0.40	-0.97	4.00	0.809
Small Metro	18.14	18.18	-0.04	1.46	5.91	0.805
Nonmetro	19.35	16.60	2.75	-1.14	2.93	0.697
Less Than High School	19.09	21.08	-1.99	1.21	8.52	0.887
High School Graduate	17.77	16.06	1.71	2.18	3.79	0.565
Some College	18.55	19.95	-1.40	-5.78	6.19	0.351
College Graduates	16.04	14.69	1.35	2.07	5.29	0.610
Employed Full Time	15.01	15.76	-0.75	0.28	3.28	0.931
Employed Part time	20.86	16.39	4.47	1.07	3.71	0.773
Unemployed	23.72	25.76	-2.04	-3.88	12.26	0.752
Others	20.55	19.92	0.63	-0.94	6.48	0.885
Health Insurance: Yes	17.00	16.42	0.58	0.63	3.01	0.834
Health Insurance: No	21.43	23.41	-1.98	-8.08	5.66	0.154
<100% Poverty Threshold	25.01	24.75	0.26	-16.72	12.29	0.174
100%-199% Poverty Threshold	22.39	21.28	1.11	2.84	3.73	0.446
≥200% Poverty Threshold	15.27	15.28	-0.01	2.46	2.87	0.391
Received Mental Health Treatment: Yes	53.92	52.83	1.09	-6.21	13.62	0.648
Received Mental Health Treatment: No	12.05	11.95	0.10	0.80	2.34	0.731

AMI = any mental illness; SDS = Sheehan Disability Scale; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

Bias measure is the weighted mean value of the difference between the true value of AMI and predicted value of AMI taken across the clinical respondent subsample within a particular subpopulation.

NOTE: See Table 3.2 for notes on domain definitions.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. Chapter 1 described the 2008A and 2008B samples.

6. Comparisons of SMI and AMI Prevalence Estimates Using the 2012 and 2008 Models

The previous chapters discussed how the 2008 model used in the estimation of serious mental illness (SMI) and any mental illness (AMI) prevalences was modified to come up with a revised 2012 model. To maintain the ability to compare changes over time, mental illness estimates would be re-computed for the data file and reports for 2008 to 2011 using the 2012 model. The purpose of the estimates in this chapter is to provide a summary of the cumulative impact of the model revisions on National Survey on Drug Use and Health (NSDUH) adult SMI and AMI prevalence estimates given the following changes:

1. Three new predictor variables have been added to the logistic model predicting SMI for the 2008A to 2012 sample: past year major depressive episode (MDE), past year serious suicidal thoughts, and an age variable.
2. The entire 2008A to 2012 clinical subsample was used to model SMI. This sample was much larger than the sample used in 2008.
3. The entire clinical subsample from 2008 (including both 2008A and 2008B) through 2012 has been used to model SMI and then to create SMI and other mental illness predicted values for the 2008B NSDUH sample (using the same covariates as the 2012 WHODAS model with the WHODAS predictor variable removed).
4. The weights used in modeling the clinical subsample have been revised as described in Section 2.7 to better account for coverage and nonresponse bias.

Both the 2008 and 2012 WHODAS methods for estimating mental illness prevalence use a standard cut point estimator derived from a model predicting SMI. The 2012 model was fit using the larger clinical dataset. Only the 2008A subsample of approximately 750 cases was used for the 2008 model; whereas the combined 2008A to 2012 subsample of approximately 5,000 cases was used for the 2012 model. The WHODAS models for 2008 and for 2012 assign each NSDUH respondent a probability of being diagnosed with SMI using this equation:

$$p_k = \frac{1}{1+e^{-\left(b_0 + b_{K6}\mathrm{Kc_7} + b_{WHODAS}\mathrm{Ac_0} + b_{MDE}\mathrm{AMDEYR2} + b_{SUI}\mathrm{MHSUITHK} + b_{AGE}AGE1830\right)}}. \qquad (15)$$

What differs are the values assigned to the coefficients in equation (15), which are zero for some coefficients in the 2008 WHODAS model.

Table 6.1 displays the coefficients for the two models as well as the cut point used to generate SMI and AMI predicted values under those models. It also displays the analogous values for the 2012 model for the Sheehan Disability Scale (SDS) sample. It does *not* display the 2008 SDS model for the 2008B sample, which was used only for SMI predicted values. For these, see Appendix B of Substance Abuse and Mental Health Services Administration (SAMHSA, 2009).

Table 6.1 The 2008 and 2012 Models and Their Cut Points

Coefficient	2008 WHODAS Model	2012 WHODAS Model	2012 Model for the SDS Sample
b_0	-4.74999920	-5.972664	-5.7736246
b_{K6}	0.20977232	0.0873416	0.1772067
b_{WHODAS}	0.38388395	0.3385193	0
b_{MDE}	0	1.1267330	1.6428623
b_{SUI}	0	1.9552664	1.8392433
b_{AGE}	0	0.1059137	0.1231266
SMI Cut Point	0.26971946	0.260573529	0.236434
AMI Cut Point	0.02400	0.0192519810	0.019182625

WHODAS = World Health Organization Disability Assessment Schedule; SDS = Sheehan Disability Scale; SMI = serious mental illness; AMI = any mental illness; MHSS = Mental Health Surveillance Study.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; MHSS clinical data, 2008-2012.
The values for the 2008 WHODAS Model use only 2008A clinical data. Chapter 1 described the 2008A sample.

Estimates in Tables 6.2 to 6.4 have been computed using combined 2008 to 2012 NSDUH data counting each represented person-year equally (i.e., the NSDUH main-sample weights in a year have *not* been scaled[37]). Instead, the prevalence estimates from each year were implicitly weighted by the number of adults in that year.[38,39] This is how prevalence estimates across years for NSDUH estimates are generally reported (e.g., see SAMHSA, 2012).

Unlike previous tables in this report, only one decimal place is displayed on subsequent tables in this chapter. That conforms to how mental illness prevalence estimates are actually published. Moreover, except when otherwise noted (AMI estimates based on the 2008 model), the estimates computed for this chapter do not use 2008B sample data.

Table 6.2 presents the estimates of SMI for the overall population and within the domains studied in this effort. As shown in the table, the 2012 model SMI estimates for the adult population and by domain were smaller than their corresponding 2008 model counterparts, especially for the 18 to 25 age group, which had a large apparent bias under the 2008 model. Furthermore, most of the 2012 model estimates were closer to the direct estimates from the Mental Health Surveillance Survey (MHSS) clinical subsample (assumed to be unbiased

[37] Scaling the annual clinical subsamples reduced the variance of estimates produced from them. Scaling the NSDUH main-survey weights analogously render measures of bias and root mean square error (RMSE) cleaner in that they are strictly because of the estimation method and not differences in how the weights are scaled. When computing prevalence estimates from the NSDUH main-survey sample for public release, however, there is no reason to scale the NSDUH main-survey weights.

[38] This means that the prevalence estimate for each year was implicitly factored by the adult population size in that year (see Section 2.7). As a result the estimate from 2012 "counted" slightly more than the estimate from 2008.

[39] The estimates from the 2008A and 2008B samples were computed using half-sample NSDUH main-survey weights and then implicitly weighted by half the number of adults in 2008 with the following exception. For AMI estimates from the 2008 model, only the 2008A sample is used. That sample's estimates are then implicitly weighted to the adult population in 2008.

estimates) than the current estimates.[40] Analogous annual estimates are displayed in Tables A.3 to A.9 in Appendix A.

Table 6.2 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models by Domain

Domain	Direct Estimate from MHSS Clinical Subsample (%)	Standard Cut Point Estimate Using 2008 Models (%)	Standard Cut Point Estimate Using 2012 Models (%)
Overall	3.9	4.9[a]	3.9
Age			
18-25	3.7	7.7[a]	3.8
26-34	4.2	6.5[a]	5.0
35-49	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
North Central	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			
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
Others	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 Graduates	2.9	3.5	3.1

(continued)

[40] The direct estimates *have been scaled* to increase statistical efficiency, with the 2008A and 2008B samples each getting a scaling factor of .06. An *F* based on the clinical subsample revealed no significant differences across the years at the .01 level in either SMI or AMI prevalence estimates, suggesting that any biasing from this scaling would be small.

Table 6.2 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models by Domain (continued)

Domain	Direct Estimate from MHSS Clinical Subsample (%)	Standard Cut Point Estimate Using 2008 Models (%)	Standard Cut Point Estimate Using 2012 Models (%)
Poverty			
<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[aa]	3.0
Health Insurance			
Yes	3.6	4.6[a]	3.6
No	5.8	6.7	5.2

SMI = serious mental illness; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

[a] The difference between the model-based and direct estimates computed from the clinical subsample (the bias measure in equation (9) of Section 2.2) is significantly different from zero at the .05 level.

NOTE: See Table 3.2 for notes on domain definitions.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. For the direct estimator, the scaling factors were described in Section 2.7.

Table 6.3 presents the estimates of AMI for the overall population and within the same domains as Table 6.2.[41] Almost all the 2012 model AMI estimates were smaller than their 2008-model counterparts, except for estimates among adults aged 50 or older. Moreover, most of the 2012 model estimates were closer to the direct estimates from the MHSS clinical subsample than the current estimates, especially the ones for the younger age groups. Analogous annual estimates are displayed in Tables A.10 to A.16 in Appendix A.

Table 6.4 compares estimates of AMI and SMI by detailed age bands. The table clearly demonstrates the differences between the estimates from the two models for the two younger age bands. For example, the SMI prevalence estimate for the 18 to 25 age band is 7.7 percent when computed from the 2008 model, but only 3.8 percent when computed from the 2012 model.

The "naïve standard errors" for SMI and AMI displayed in this table were computed by treating the predicted SMI and AMI values as true values. They are labeled "naïve" because they fail to account for the error from modeling SMI. These standard-error measures are nonetheless useful when comparing the estimates from different age bands because the standard error of a difference is roughly equal to the square root of the sum of the squares of the individual naïve standard errors. As shown in Table 6.4, most of the naïve standard errors based on the 2012 model are equal to or smaller than their counterparts based on the 2008 model.

[41] AMI estimates computed with the 2008 model do not use data from the 2008B sample except in Appendix A.

Table 6.3 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models by Domain

Domain	Direct Estimate from MHSS Clinical Subsample (%)	Standard Cut Point Estimate Using 2008 Model (%)	Standard Cut Point Estimate Using 2012 Models (%)
Overall	18.0	20.0	18.1
Age			
18-25	20.9	30.5[a]	18.5
26-34	19.7	24.1	22.0
35-49	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
North Central	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			
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
Others	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 Graduates	16.7	17.0	16.1
Poverty			
<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; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

[a] The difference between the model-based and direct estimates computed from the clinical subsample (the bias measure in equation (9) of Section 2.2) is significantly different from zero at the .05 level.

NOTE: See Table 3.2 for notes on domain definitions.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. The 2008B sample was not used in 2008 model estimates. For the direct estimator, the scaling factors were described in Section 2.7. Chapter 1 described the 2008A and 2008B samples.

Table 6.4 Comparing Mental Health Prevalence Estimates Using the 2008 and 2012 Models, by Detailed Age Group

Age Category	AMI (2008 Model)		AMI (2012 Models)		SMI (2008 Models)		SMI (2012 Models)	
	Estimate (%)	Naïve SE (%)	Estimate (%)	Naïve SE (%)	Estimate (%)	Naïve SE (%)	Estimate (%)	Naïve SE (%)
TOTAL	20.0	0.15	18.1	0.13	4.9	0.07	3.9	0.07
18	32.1	0.56	16.8	0.40	7.7	0.29	3.3	0.19
19	32.2	0.56	17.3	0.42	7.6	0.29	3.4	0.21
20	31.9	0.57	18.1	0.43	8.6	0.32	3.8	0.22
21	31.1	0.56	18.0	0.43	7.9	0.29	3.9	0.23
22	29.7	0.55	18.3	0.43	7.2	0.29	3.6	0.21
23	28.9	0.57	19.4	0.45	7.8	0.30	4.4	0.21
24	28.9	0.55	19.7	0.46	7.4	0.29	3.9	0.21
25	28.3	0.61	21.3	0.47	7.1	0.30	4.1	0.22
26-29	25.5	0.51	21.5	0.45	6.6	0.27	4.8	0.23
30-34	22.9	0.47	22.4	0.43	6.4	0.25	5.2	0.22
35-39	21.9	0.47	21.4	0.42	6.0	0.25	5.2	0.24
40-44	20.0	0.45	20.0	0.41	5.2	0.22	4.8	0.21
45-49	19.8	0.43	19.7	0.42	5.3	0.22	5.0	0.22
50-54	17.4	0.58	17.5	0.52	4.8	0.26	4.4	0.25
55-59	16.7	0.63	17.5	0.58	4.0	0.29	3.6	0.27
60-64	14.7	0.63	15.6	0.60	3.1	0.28	3.3	0.30
65 or Older	11.3	0.38	12.0	0.35	1.4	0.12	1.2	0.12

SMI = serious mental illness; AMI = any mental illness; SE = standard error; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study. *Naïve* SE = standard error computed as if predicted AMI (SMI) values were actual values.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. For the direct estimator, the scaling factors were described in Section 2.7. Chapter 1 described the 2008A and 2008B samples. Only the 2008A sample is used to compute the AMI estimates using the 2008 model.

7. Some Concluding Comments

In this chapter, we conclude with several comments and caveats:

- We remind the reader that the target population of the National Survey on Drug Use and Health (NSDUH) includes adults living in households, including noninstitutional group quarters.
- We explore why the NSDUH-based estimates of serious mental illness (SMI) and any mental illness (AMI) are lower than other published estimates.
- We discuss how the error rate, a focus in model creation in this report, relates to sensitivity and specificity.
- We provide caution about using the model-derived NSDUH estimates of SMI (or AMI) in models containing variables related to suicidal ideation or the experience of a major depressive episode (MDE).

7.1 The Target Population

The target population for the annual NSDUH is 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). Although the sampling frame does not directly include persons who had no fixed address (e.g., homeless and/or transient persons not in shelters), who were on active military duty, who lived in institutional group quarters, or who resided in institutional group quarters (e.g., correctional facilities, nursing homes, mental institutions, and long-term hospitals), it does include persons who spent at least some time during the year in a household, homeless shelter, or other group quarters.

Using Census Bureau data, we estimate the target population did not include roughly 5 million of 240 million (around 2%) in the adult population living in the United States in 2012. Of those 5 million, 20% (1 million) were in the armed forces. Even if the SMI prevalence among the nonmilitary population not covered by the NSDUH was twice as high as the NSDUH estimate, the increase in the national prevalence estimate would be less than 0.5%.

7.2 Comparability of NSDUH's SMI and AMI Prevalence Estimates with Other National Estimates

The National Comorbidity Survey (NCS) and National Comorbidity Survey Replication (NCS-R) (Harvard School of Medicine, 2005) also contain estimates of past-year SMI and AMI among adults residing in U.S. households. The former is based on interviews in 1990 and 1991 of English-speaking adults 18 to 54, and the latter on interviews between 2001 and 2003 of English-speaking adults of all ages. Like the NSDUH, adults not living in households were excluded from the target population.

The NCS and NCS-R estimates of past-year SMI among adults were considerably higher than those from the NSDUH. The NCS-R estimate for the English-speaking adult population in

households was 5.8%. The NCS estimate, when adjusted for the absence of adults aged 55 or older using data from the prior Epidemiological Catchment Area (ECA) study data from one site (Baltimore), was comparable to the NCS-R estimate (5.4%). One source of difference is which disorders were included in the definitions. For example, the NSDUH definition and estimate of both SMI and AMI excludes substance use disorders, while in most publications, the NCS-R estimate does not. When substance use disorder is added, a NSDUH AMI estimate of 22.5% was close to an NCS-R estimate of 26.2%.[42]

Several factors may contribute to the higher SMI (and AMI) prevalence estimates for the NCS-R. One small contributor to the difference between the NCS-R and NSDUH estimates is that the NCS-R weighting procedures did not adjust for missing Spanish speakers who could not (or chose not to) respond in English. Data from the NSDUH suggests this group has a lower SMI prevalence compared to all adults in the United States. Therefore, adjusting for their exclusion (as in the NSDUH/MHSS) will slightly lower the overall prevalence of SMI. A more significant contributor most likely is that the NCS-R estimates do not appear to have adjusted for the fact that respondents tend to have a relatively higher probability of having SMI than nonrespondents. This was determined from a nonresponse bias analysis that compared NSDUH main study mental health-related responses for both respondents and nonrespondents to the clinical interview (Chapter 5 of Center for Behavioral Health Statistics and Quality [CBHSQ], 2014). NSDUH estimates were adjusted to remove this potential source of response bias. Documentation available on the NCS-R estimation methods does not indicate that the estimates were adjusted analogously.

Although significant effort was made to measure and adjust for the potential selection biases among NSDUH adults who responded to the clinical interview (CBHSQ, 2014), there is no guarantee that the adjustments for nonresponse to the clinical interview and the undercoverage of Hispanics who chose to respond to the main survey in Spanish actually removed 100% of all such bias. Similarly, there is no guarantee that the nonresponse and coverage adjustments in the full NSDUH or the NCS-R were completely successful in removing all of the biases.

Differences in survey mode, instrumentation, interviewers, and specific measures used to assess different mental disorders may also contribute to the differences in the prevalence estimates derived from these two data sources. Although both surveys used DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders*, 4th ed.; American Psychiatric Association, 1994) criteria to determine whether respondents had selected mental health disorders in the past year differed. The MHSS follow-up study used trained clinical interviewers to administer a semistructured telephone interview based on the Structured Clinical Interview for DSM-IV (SCID; as discussed in Chapter 1), whereas the NCS and NCS-R used lay interviewers to administer a structured interview in person, based on the Composite International Diagnostic Interview (CIDI; Robins, Wing, Wittchen, & Helzer, 1988). The SCID is considered the clinical gold standard in determining mental disorder status, and is the instrument against which the CIDI is validated.

[42] From Table B.1 of Karg and colleagues (2014). Appendix B of that report supplies a more thorough description of the differences between NSDUH and NCS-R results than contained here.

More crucially for SMI, the NSDUH clinicians made assessments of functional impairment using the Global Assessment of Functioning (GAF) score from the SCID structured interview used to diagnose psychological disorders. Respondents with a GAF score of 50 or below and a psychological disorder were assessed to have past year SMI. In contrast to this, in the NCS-R, severe functional impairment was determined based on responses to the CIDI-based interview using this definition:

> [past-year] cases were classified serious if they had any of the following: a 12-month suicide attempt with serious lethality intent; work disability or substantial limitation due to a mental or substance disorder; a positive screen for non-affective psychosis; bipolar I or II disorder; substance dependence with serious role impairment (as defined by disorder-specific impairment questions); an impulse-control disorder with repeated serious violence; or any disorder that resulted in 30+ days out of role in the year (Kessler, Chiu, Demler, & Walters, 2005, p. 619).

The NSDUH clinical interviewers were much less likely to judge the level of functional impairment exhibited in their SCID-based interviews sufficient to warrant an SMI diagnosis. Only an estimated 4.5% of adults had GAF scores of 50 or below, while an estimated 5.8% have some mental disorder and a GAF of 53 or below.

Context effects also may have had an effect on respondents' answers in the two studies. The order in which mental disorders were assessed in the two studies was not identical. Context effects occur when prior questions affect responses to later questions in surveys. A respondent may answer a subsequent question in a manner that is consistent with responses to a preceding question if the two questions are closely related to each other (e.g., a respondent denies use of both cocaine and amphetamines because they are both in the stimulant drug class). As an example, the 2008 NSDUH found that the inclusion of new items to assess functional impairment and suicidality before the questions on depression altered the estimates of adult MDE relative to previous years, even though the depression questions themselves did not change (Office of Applied Studies, 2009).

The order in which disorders are assessed in the CIDI (NCS-R) and the SCID (Mental Health Surveillance Survey [MHSS] clinical study) differed. For example, in the MHSS clinical study, psychotic symptoms and posttraumatic stress disorder (PTSD) were assessed near the beginning of the SCID (following the assessment of mood disorders), but in the NCS-R, they were assessed near the end of the CIDI (following the assessments of mood disorders, other anxiety disorders, substance use disorders, eating disorders, and impulse control disorders). In addition, the MHSS clinical study assessed lifetime MDE and manic episode only if past year disorder was absent to enable proper past year diagnosis of a major depressive disorder (MDD) versus bipolar disorder; however, the MHSS did not assess lifetime occurrence for any other disorders. In contrast, the NCS-R assessed lifetime occurrence for all disorders, followed by an assessment of past year disorder only if the disorder was present in the person's lifetime. These differences may explain why some prevalence past year estimates are higher in the NCS-R than in the MHSS clinical study.

In addition, the MHSS clinical study and the NCS-R each used screening questions for certain disorders to determine whether further assessment was warranted. The SCID used in the

MHSS clinical study included a section of screener questions at the beginning of the interview for each of the following disorders: panic disorder, agoraphobia, social phobia, specific phobia, obsessive compulsive disorder (OCD) (two questions—one for obsessions and one for compulsions), general anxiety disorder (GAD), anorexia nervosa, and bulimia nervosa. The SCID also included screener questions at the beginning of the PTSD, mood episodes, and intermittent explosive disorder modules, where a "no" answer to a particular question skips the respondent out of the remainder of that module. The CIDI used in the NCS-R included *multiple* screening questions at the beginning of the interview. These questions included screeners for depression (MDD; three questions), mania (bipolar disorder; three questions), panic disorder (two questions including one "second chance" question to query physical symptoms that come on suddenly but are not attributed to fear or panic), social phobia (four questions), agoraphobia (four questions), GAD (three questions), intermittent explosive disorder (three questions), and specific phobia (six questions). A respondent was then routed into a disorder-specific module after the completion of the screening questions if at least one screening question endorsed. These differences in interview context and screening patterns may explain some of the differences in prevalence estimates across surveys. Karg et al. (2014, Appendix B) provides a detailed description of the methods used in the MHSS clinical study and the NCS-R that may contribute to differences in estimates between the two studies.

7.3 The Sensitivity and Specificity of SMI and AMI Prevalence Estimates

This report focuses on the error rate as a metric to assess the 2012 prediction model. In this section, the sensitivity and specificity of alternative models are evaluated because they are also commonly used metrics. Table 7.1 displays the sensitivity and specificity of the same SMI models appearing in Table 5.2. Paralleling Table 5.2, Table 7.1 displays the properties of the AMI all-adult prevalence estimates in addition to SMI prevalence estimates because the SMI model is used to produce both estimates.

The sensitivity of an SMI model in producing a prevalence estimator for SMI (or AMI) is the estimated fraction of adults with SMI (AMI) according to the clinical interview who are assigned SMI (AMI) status by the model. The specificity of a model is the estimated fraction of adults without SMI (AMI) who are not assigned SMI (AMI) status. The former is relatively small (and the latter relatively large) because the fraction of adults estimated to have (or not have) SMI is small (or large). Mathematically,

$$Sensitivity = 1 - \frac{\text{False negative rate}}{\text{Prevalence rate}} \Rightarrow \textit{False negative rate} = \text{Prevalence rate} \times (1 - \text{Sensitivity}),$$

and

$$Specificity = 1 - \frac{\text{False positive rate}}{100\% - \text{Prevalence rate}} \Rightarrow \textit{False positive rate} = (100\% - \text{Prevalence rate}) \times (1 - \text{Specificity}).$$

Table 7.1 Sensitivity and Specificity of Alternative Models for SMI in Producing SMI and AMI All-Adult Prevalence Estimates

Predictor Variables	Error Rate (%)	False Positive Rate (%)	False Negative Rate (%)	Sensitivity	Specificity
For SMI					
Kc_7	5.31	2.75	2.55	0.351	0.971
Kc_7 Ac_0	4.42	2.21	2.21	0.438	0.977
Kc_7 Ac_0 AGE18 AGE30	4.30	2.16	2.14	0.456	0.978
Kc_7 Ac_0 AGE1830	4.30	2.17	2.12	0.460	0.977
Kc_7 Ac_0 MHSUITHK AMDEYR2	3.97	1.99	1.98	0.497	0979
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	3.76	1.88	1.88	0.522	0.980
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830[a]	3.84	1.92	1.93	0.509	0.980
For AMI					
Kc_7	17.43	8.43	8.99	0.501	0.897
Kc_7 Ac_0	15.85	7.99	7.86	0.564	0.903
Kc_7 Ac_0 AGE18 AGE30	15.50	7.75	7.75	0.570	0.851
Kc_7 Ac_0 AGE1830	15.58	7.79	7.79	0.568	0.828
Kc_7 Ac_0 MHSUITHK AMDEYR2	15.77	7.92	7.85	0.565	0.903
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE18 AGE30	15.62	7.78	7.84	0.565	0.905
Kc_7 Ac_0 MHSUITHK AMDEYR2 AGE1830[a]	15.47	7.70	7.77	0.569	0.906

See Chapter 4 for variable definitions.

[a] This is the model used operationally.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008A-2012 MHSS clinical respondents. The scaled weights used in combining data across years were described Section 2.7. Chapter 1 described the 2008A sample.

Observe that the sensitivity increases as the false negative rate decreases, while the specificity increases as the false positive rate decreases. A small estimated prevalence will, in most cases, produce a small sensitivity and a large specificity.

7.4 Limitations in Using the Mental Illness Variables in Analyses

There are many advantages to using the cut point methodology described in the previous chapter to predict SMI and AMI status to every adult responding to the NSDUH main-survey interview. There are, however, some analyses for which these predicted values should not be used. In particular, they should not be used in analyses that use the mental illness variables in conjunction with variables used, or closely related to those used, in the prediction model.

Table 7.2 displays how SMI and AMI prevalence estimates from the 2008 and 2012 models differed within certain domains defined by predictor variables included in the 2012 model but not in the 2008 model. These estimates are whether the adult had thoughts of suicide or had an MDE in the previous year from the NSDUH main survey.[43] Other domains displayed

[43] In creating the SMI and AMI predicted values under the 2012 models, missing values for past year suicidal thoughts or MDE were treated as zeroes. For creating the domains in Table 7.5, they are treated as missing (i.e., not allocated to a domain).

in the table have a close definitional relationship to the two predictor variables: whether the adult had suicidal plans or attempts in the past year and whether the adult had an MDE in his/her lifetime.

Table 7.2 Comparing SMI and AMI Prevalence Estimates for MDE and Suicidal Thoughts Domains Using the 2008 and 2012 Models

Domain	Direct Estimate from MHSS Clinical Subsample (%)	Standard Cut Point Estimator Using 2008 Models (%)	Standard Cut Point Estimator Using 2012 Models (%)
SMI			
Past Year MDE			
Yes	33.0	44.6	48.9[a]
No	1.8	2.0[a]	0.7[a]
Lifetime MDE			
Yes	20.5	28.0	27.0[a]
No	1.4	1.5[a]	0.5[a]
Had Suicidal Thoughts in the Past Year			
Yes	38.9	40.5[a]	57.0[a]
No	2.6	3.5	1.8[a]
Made Suicidal Plans in the Past Year			
Yes	56.1	53.6[a]	66.6[a]
No	3.4	4.4	3.2
Had Suicidal Attempt in the Past Year			
Yes	47.1	49.7	59.1
No	3.8	4.7	3.6
AMI			
Past Year MDE			
Yes	72.1	87.2[a]	98.8[a]
No	14.0	15.1[a]	12.2
Lifetime MDE			
Yes	51.5	65.8[a]	69.5[a]
No	12.9	13.1[a]	10.4[a]
Had Suicidal Thoughts in the Past Year			
Yes	65.7	78.7[a]	99.6[a]
No	16.1	17.6	14.9
Made Suicidal Plans in the Past Year			
Yes	77.7	84.5	99.6[a]
No	17.3	19.2	17.2

(continued)

Table 7.2 Comparing SMI and AMI Prevalence Estimates for MDE and Suicidal Thoughts Domains Using the 2008 and 2012 Models (continued)

Domain	Direct Estimate from MHSS Clinical Subsample (%)	Standard Cut Point Estimator Using 2008 Models (%)	Standard Cut Point Estimator Using 2012 Models (%)
Had Suicidal Attempt in the Past Year			
Yes	90.2	81.9[a]	99.5[a]
No	17.7	19.6	17.6

SMI = serious mental illness; AMI = any mental illness; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health.

[a] The difference between the model-based and direct estimates computed from the clinical subsample (the bias measure in equation (9) of Section 2.2) is significantly different from zero at the .05 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. The 2008B sample was not used for 2008 model estimates of AMI. For the direct estimator, the scaling factors were described in Section 2.7. Chapter 1 described the 2008A and 2008B samples.

Table 7.2 shows that SMI and AMI cut point estimates for these domains were often not very close to the direct estimates from the MHSS clinical subsample. Estimates using the 2012 model, especially, appear to overestimate the proportions of adults with past year suicidal thoughts and MDE having SMI or AMI.

This overestimation is likely a consequence of using a cut point methodology to estimate mental illness prevalence in the NSDUH main-survey sample. When a variable is added to a logistic model for SMI because of its large |t| value, there is a tendency for the resulting cut point predicted values for SMI (and AMI) to exhibit a larger correlation with that variable than exists between the variable and the clinical diagnosis of SMI. For that reason, care was taken to make sure that did *not* happen when AGE1830 was added to the model as described in Chapter 6.[44]

The tendency for estimates based on the 2012 model to be biased upward also existed in domains having definitions overlapping past year suicidal thoughts and MDE, like past year suicidal *plans* and lifetime MDE (e.g., an adult with a suicidal plan also had a suicidal thought, while an adult with a past year MDE also had a lifetime MDE). Similarly, but less strongly, using the 2012 model tended to underestimate the proportion of adults with SMI or AMI who did not have past year suicidal thoughts or MDE. Clearly, the mental illness prevalence estimates derived from 2012 model should not be used for these domains or for domains with overlapping definitions.

Tables 7.3 and 7.4 show that *both* the 2008 and 2012 models often produce significantly biased SMI (Table 7.3) and AMI (Table 7.4) estimates within domains defined by Kessler 6 (K6) and World Health Organization Assessment Schedule (WHODAS) scores. Model-based mental illness prevalence estimates should not be used for domains defined by the components of the K6 and WHODAS scores.

[44] The decision to base the standard cut point estimator in the 2012 method on the model fit for SMI, rather than the fit for AMI or the joint fit was largely the result of the smaller bias measures resulting from using the SMI fit.

Table 7.3 Comparing SMI Prevalence Estimates for K6 and WHODAS Domains Using the 2008 and 2012 Models

	Direct Estimate from MHSS Clinical Subsample (%)	Standard Cut Point Estimator Using 2008 Models (%)	Standard Cut Point Estimator Using 2012 Models (%)
Worse K6 Total Score in Past Year (KSCMAX6)			
0-3	0.4	0.0[a]	0.0[a]
4-5	0.5	0.0	0.1
6-7	2.7	0.0[a]	0.3[a]
8-9	5.5	0.0[a]	1.0[a]
10-11	8.0	1.4[a]	3.0[a]
12-15	11.7	12.9	11.2
16-24	29.7	62.8[a]	45.7[a]
Alternative WHODAS Total Score (WHODASC3)			
0	0.7	0.0[a]	0.1[a]
1	2.7	0.4[a]	0.9
2	5.6	1.5[a]	2.3
3	8.2	3.9[a]	4.8
4	11.2	13.7[a]	9.3
5	22.2	30.2	19.1
6	30.3	52.2	35.1
7	31.9	65.0[a]	50.1[a]
8	39.6	86.6[a]	71.6[a]

K6 = Kessler 6; NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

[a] The difference between the model-based and direct estimates computed from the clinical subsample with scaled weights (the bias measure in equation (9) of Section 2.2) is significantly different from zero at the .05 level.

NOTE: See Section 4.1 for the definition of WHODASC3.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. For the direct estimator, the scaling factors were described in Section 2.7.

Table 7.4 Comparing AMI Prevalence Estimates for K6 and WHODAS Domains Using the 2008 and 2012 Models

	Direct Estimate from MHSS Clinical Subsample (%)	Standard Cut Point Estimator Using 2008 Models (%)	Standard Cut Point Estimator Using 2012 Models (%)
Worse K6 Total Score in Past Year (KSCMAX6)			
0-3	5.2	1.0[a]	1.5[a]
4-5	14.8	6.2[a]	6.9[a]
6-7	18.8	14.3[a]	15.3
8-9	34.9	28.6[a]	28.1
10-11	33.8	55.9	50.6[a]
12-15	47.6	100.0[a]	73.4[a]
16-24	69.3	100.0	95.0[a]
Alternative WHODAS Total Score (WHODASC3)			
0	8.9	3.5[a]	2.8[a]
1	18.8	23.1	18.3
2	37.0	44.9	40.3
3	40.4	100.0[a]	84.6[a]
4	50.6	100.0[a]	92.5[a]
5	62.0	100.0	97.8[a]
6	69.7	100.0[a]	99.8[a]
7	71.5	100.0	100.0[a]
8	76.5	100.0[a]	100.0[a]

AMI = any mental illness; K6 = Kessler 6; NSDUH = National Survey on Drug Use and Health; WHODAS = World Health Organization Disability Assessment Scale.

[a] The difference between the model-based and direct estimates computed from the clinical subsample with scaled weights (the bias measure in equation (9) of Section 2.2) is significantly different from zero at the .05 level.

NOTE: See Section 4.1 for the definition of WHODASC3.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. The 2008B sample was not used for 2008 model estimates. For the direct estimator, the scaling factors were described in Section 2.7. Chapter 1 described the 2008A and 2008B samples.

Analysts are sometimes interested not only in the SMI prevalence within different domains, but in the attributes of adults within domains defined by having (or not having) SMI. The final two tables look at the suicidal plans and attempts in such domains. Table 7.5 displays the proportions of suicidal plans and attempts among adults with SMI as estimated by the clinical subsample using the clinical determination of SMI and as estimated in the NSDUH main-survey sample using the 2008 and 2012 models, respectively to predict SMI. Both model-derived estimates were biased for plans, neither for attempts, with the 2012 estimates higher in both cases.

Table 7.6 displays the proportions of adults who planned or attempted suicide among those who had suicidal thoughts estimated within the two SMI domains (has/does not have) defined by either the clinical evaluation or one of the models. Here the proportions for the domain with SMI were less when the 2012 model was used than the 2008 model. Our ability to detect statistically significant differences was hampered by the relatively small number of adults

with suicidal thoughts in the clinical subsample (288 with clinically determined SMI, 326 without).

Table 7.5 Comparing Proportions Planning or Attempting Suicide in the Past Year Among Adults with SMI Using the 2008 and 2012 Models

Estimate	Direct Estimate from MHSS Clinical Sample (%)	Estimate from NSDUH Main Sample Using 2008 SMI Predicted Values (%)	Estimate from NSDUH Main Sample Using 2012 SMI Predicted Values (%)
Proportion with SMI who Planned Suicide	14.9	11.7[a]	18.3[a]
Proportion with SMI who Attempted Suicide	4.3	5.0	7.6

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

[a] The difference between the model-based and direct estimates computed from the clinical subsample (the bias measure in equation (9) of Section 2.2 extended to ratio estimates by repeating each observation, assigning one copy to the clinical SMI determination and the other the model prediction) is significantly different from zero at the .05 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. For the direct estimator, the scaling factors were described in Section 3.7.

The results in Tables 7.2 through 7.6 demonstrate the potential dangers of using the predicted SMI variable from the NSDUH main-survey sample in analyses also containing predictor variables used in its development (like K6 components and thoughts of suicide). Recall, however, that replacing a 0/1 predicted value for SMI with an estimated probability of having SMI is not a viable alternative as the many domain-level biases in Table 3.2 attest.

Table 7.6 Comparing Proportions Planning or Attempting Suicide in the Past Year Among Adults with Suicidal Thoughts by SMI Domain Using the 2008 and 2012 Models

Domain	Direct Estimate from MHSS Clinical Sample (%)	Estimate from NSDUH Main Sample Using 2008 SMI Predicted Values (%)	Estimate from NSDUH Main Sample Using 2012 SMI Predicted Values (%)
Planned Suicide			
Had SMI	40.5	37.6	33.2[a]
Did not have SMI	20.0	22.1	22.0
Attempted Suicide			
Had SMI	11.5	16.2	13.7[a]
Did not have SMI	8.2	11.2	12.5[a]

MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SMI = serious mental illness.

[a] The difference between the model-based and direct estimates computed from the clinical subsample (the bias measure in equation (9) of Section 2.2 extended to ratio estimates by repeating each observation, assigning one copy to the clinical SMI determination and the other the model prediction) is significantly different from zero at the .05 level.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. The weights were not scaled for the cut point estimators when combining data across years. For the direct estimator, the scaling factors were described in Section 2.7.

References

Aldworth, J., Colpe, L., Gfroerer, J., Novak, S, Chromy, J., Barker P., Barnett-Walker, K., Karg, R., Morton, K., & Spagnola, K (2010a). The National Survey on Drug Use and Health and Mental Health Surveillance Study: Calibration analysis, *International Journal of Methods in Psychiatric Research, 19*(S1), 61-87.

Aldworth, J., Barnett-Walker, K., Chromy, J., Karg, R., Kott, P., & Morton, K. (2010b). *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, Deliverable No. 7, RTI/0209009.523.006.008). 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 Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th Ed., Text Revision). Washington, DC: Author.

Center for Behavioral Health Statistics and Quality. (2012). *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. (2013). *Results from the 2012 National Survey on Drug Use and Health: Mental health findings* (HHS Publication No. SMA 13-4805, NSDUH Series H-39). Rockville, MD: Substance Abuse and Mental Health Services Administration.

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.

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.

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: Biometrics Research, New York State Psychiatric Institute.

Harvard School of Medicine. (2005). *National Comorbidity Survey (NCS) and National Comorbidity Survey Replication (NCS-R)*. Retrieved from http://www.hcp.med.harvard.edu/ncs/

Judkins, D. R. (1990). Fay's method for variance estimation. *Journal of Official Statistics, 6*, 223-239

Karg, R. S., Bose, J., Batts, K. R., Forman-Hoffman, V. L., Liao, D., Hirsch, E., Pemberton, M. R., Colpe, L. J., & Hedden, S. L. (2014, October). *CBHSQ Data Review: Past year mental disorders among adults in the United States: Results from the 2008-2012 Mental Health Surveillance Study.* Rockville, MD: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality.

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.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 617-627.

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., Kott, P., Aldworth, J., Yu, F., Karg, R., Shook-Sa, B., 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). Retrieved from http://www.samhsa.gov/data/sites/default/files/NSDUH2011MRB/NSDUH2011MRB/NSDUHmrbMHSSDesignEst2011.pdf

McCullagh P (1980). Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society*, Series B, *42*(2), 109-142.

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. (2005). Results from the 2004 National Survey on Drug Use and Health: National findings. (NSDUH Series H-28, DHHS Publication No. SMA 05-4062). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

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

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*, 110-123.

Robins, L. N., Wing, J., Wittchen, H. U., & Helzer, J. E. (1988). The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Archives of General Psychiatry *45*, 1069-1077.

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

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

Wolter, K. M. (2007). *Introduction to Variance Estimation (*2nd Ed.). New York: Springer.

This page intentionally left blank

Appendix A: Comparisons of Estimates Using the 2008 and 2012 Models

This page intentionally left blank

Table A.1 Comparing Standard Cut Point Estimates for the SMI 2008B Sample Using a 2012 Model for that Sample Containing an SDS Variable in Place of Ac_0 with Estimates from the 2008A Sample by Domain

Domain	Standard Cut Point Estimates (%)			Bias Measures for 2008B Estimates (%)		
	2008A	2008B	Difference	Measure	Standard Error	*p*-value
All Adults	3.72	3.71	-0.01	-2.54	1.80	0.158
Age: 18-25	4.15	2.62	-1.53	-1.15	0.85	0.175
Age: 26-34	4.57	4.88	0.31	-0.69	2.16	0.750
Age: 35-49	4.74	4.84	0.10	-0.91	1.62	0.572
Age: 50 or Older	2.52	2.86	0.34	-4.90	4.14	0.236
Health Insurance: Yes	3.54	3.43	-0.11	-1.91	1.89	0.311
Health Insurance: No	4.71	5.31	0.60	-9.05	5.80	0.119
< 100% of the Poverty Threshold	6.25	6.11	-0.14	-17.81	11.33	0.116
100%-199% of the Poverty Threshold	4.84	4.98	0.14	-0.37	1.24	0.767
≥ 200% the Poverty Threshold	3.02	3.00	-0.02	-0.07	0.64	0.915
White, Not Hispanic	4.14	4.25	0.11	-2.42	2.47	0.326
Black, Not Hispanic	2.66	3.01	0.35	-3.01	4.37	0.491
Other, Not Hispanic	3.40	1.93	-1.47	-7.82	5.23	0.135
Hispanic	2.58	2.42	-0.16	-0.24	0.25	0.337
Male	2.76	2.35	-0.41	-1.94	1.29	0.134
Female	4.61	4.98	0.37	-3.10	3.24	0.339
Northeast	3.36	4.05	0.69	-5.15	5.49	0.348
North Central	4.23	3.65	-0.58	0.41	0.64	0.525
South	3.71	3.46	-0.25	-2.88	2.10	0.170
West	3.53	3.90	0.37	-1.46	2.24	0.514
Large Metro	3.55	3.53	-0.02	0.06	0.77	0.941
Small Metro	3.82	4.09	0.27	-4.52	4.58	0.324
Nonmetro	4.07	3.61	-0.46	-4.46	3.05	0.143
Received Mental Health Treatment: Yes	18.39	19.41	1.02	-16.87	11.32	0.136
Received Mental Health Treatment: No	1.43	1.28	-0.15	-0.32	0.47	0.504
Employed Full Time	2.69	3.00	0.31	-1.05	1.04	0.315
Employed Part Time	4.06	2.97	-1.09	-1.56	1.79	0.385
Unemployed	7.98	5.40	-2.58	6.05	4.31	0.161
Other Employment Status	5.00	5.17	0.17	-5.99	5.07	0.237
Less Than High School	3.28	3.74	0.46	-9.75	7.11	0.170
High School Graduate	4.38	3.61	-0.77	0.52	0.62	0.409
Some College	4.33	4.73	0.40	-5.75	5.48	0.294
College Graduate	2.65	2.89	0.24	-0.95	1.18	0.421

SMI = serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study; SDS = Sheehan Disability Scale.
NOTE: See Table 3.2 for notes on domain definitions.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. Chapter 1 described the 2008A and 2008B samples.

Table A.2 Comparing Standard Cut Point Estimates for the AMI 2008B Sample Using a 2012 Model for that Sample Containing an SDS Variable in Place of Ac_0 with Estimates from the 2008A Sample by Domain

	Standard Cut Point Estimates (%)			Bias Measures for 2008B Estimates (%)		
Domain	**2008A**	**2008B**	**Difference**	**Measure**	**Standard Error**	***p*-value**
All Adults	17.69	17.46	-0.23	-0.14	2.78	0.960
Age: 18-25	18.83	16.36	-2.47	-3.87	3.32	0.244
Age: 12-34	20.75	21.68	0.93	-4.92	5.54	0.375
Age: 35-49	20.08	21.13	1.05	4.09	5.02	0.415
Age: 50 or Older	14.43	13.66	-0.77	0.53	5.31	0.920
Health Insurance: Yes	17.00	16.42	-0.58	0.63	3.01	0.834
Health Insurance: No	21.43	23.41	1.98	-8.08	5.66	0.154
< 100% of the Poverty Threshold	25.01	24.75	-0.26	-16.72	12.29	0.174
100%-199% of the Poverty Threshold	22.39	21.28	-1.11	2.84	3.73	0.446
≥ 200% of the Poverty Threshold	15.27	15.28	0.01	2.46	2.87	0.391
White, Not Hispanic	18.44	17.73	-0.71	-1.57	3.45	0.649
Black, Not Hispanic	15.79	18.07	2.28	-5.90	7.20	0.413
Other, Not Hispanic	16.70	16.55	-0.15	16.01	9.88	0.105
Hispanic	15.91	15.99	0.08	4.27	7.40	0.564
Male	13.45	13.93	0.48	4.27	3.90	0.273
Female	21.64	20.74	-0.90	-4.25	3.90	0.276
Northeast	16.68	19.97	3.29	1.14	7.55	0.880
North Central	18.64	17.64	-1.00	4.93	2.49	0.047
South	17.50	16.16	-1.34	-7.07	3.85	0.066
West	17.89	17.31	-0.58	3.24	6.88	0.637
Large Metro	16.91	17.31	0.40	-0.97	4.00	0.809
Small Metro	18.14	18.18	0.04	1.46	5.91	0.805
Nonmetro	19.35	16.60	-2.75	-1.14	2.93	0.697
Received Mental Health Treatment: Yes	53.92	52.83	-1.09	-6.21	13.62	0.648
Received Mental Health Treatment: Yes	12.05	11.95	-0.10	0.80	2.34	0.731
Employed Full Time	15.01	15.76	0.75	0.28	3.28	0.931
Employed Part Time	20.86	16.39	-4.47	1.07	3.71	0.773
Unemployed	23.72	25.76	2.04	-3.88	12.26	0.752
Other Employment Status	20.55	19.92	-0.63	-0.94	6.48	0.885
Less Than High School	19.09	21.08	1.99	1.21	8.52	0.887
High School Graduate	17.77	16.06	-1.71	2.18	3.79	0.565
Some College	18.55	19.95	1.40	-5.78	6.19	0.351
College Graduate	16.04	14.69	-1.35	2.70	5.29	0.610

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study; SDS = Sheehan Disability Scale.

NOTE: See Table 3.2 for notes on domain definitions.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality; 2008-2012 NSDUH and MHSS clinical respondents. Chapter 2 described the 2008A and 2008B samples.

Table A.3 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008A NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Model	Standard Cut Point Estimate Using 2012 Model
Overall	4.6	4.4	3.7
Age			
18-25	4.9	7.6	4.2
26-34	6.1	5.5	4.6
35-49	4.1	5.0	4.7
50 or Older	4.0	2.3	2.5
Gender			
Male	3.7	3.0	2.8
Female	5.4	5.6	4.6
Race			
White, Not Hispanic	6.1	4.9	4.1
Black, Not Hispanic	2.2	3.0	2.7
Other, Not Hispanic	0.6	3.2	3.4
Hispanic	0.6	3.3	2.6
Region			
Northeast	2.8	4.2	3.4
North Central	4.2	4.6	4.2
South	3.7	4.5	3.7
West	8.4	4.0	3.5
County Type			
Large Metro	3.9	4.1	3.6
Small Metro	5.9	4.3	3.8
Nonmetro	3.8	5.4	4.1
Received Mental Health Treatment			
Yes	18.0	19.9	18.4
No	1.6	2.0	1.4
Employment Level			
Employed Full Time	1.9	3.2	2.7
Employed Part Time	6.1	5.4	4.1
Unemployed	20.6	9.1	8.0
Others	10.3	5.5	5.0
Education			
Less than High School	10.8	4.6	3.3
High School Graduate	6.2	4.9	4.4
Some College	3.1	5.1	4.3
College Graduate	3.1	3.0	2.7
Poverty			
<100% Threshold	6.4	7.4	6.2
100%-199% Threshold	12.9	6.0	4.8
≥200% Threshold	3.0	3.4	3.0
Health Insurance			
Yes	3.9	4.1	3.5
No	7.6	5.9	4.7

SMI =serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTES: In 2008, a split-sample design assigned adults aged 18 or older randomly to one of two impairment scales, the World Health Organization Disability Assessment Schedule (WHODAS) or the Sheehan Disability Scale (SDS). Estimates here are based on the WHODAS half-sample (2008A). See Table 3.2 for notes on domain definitions.

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

Table A.4 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008B NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Model	Standard Cut Point Estimate Using 2012 Model
Overall	4.8	4.4	3.7
Age			
18-25	4.1	7.5	3.4
26-34	4.4	6.0	4.8
35-49	3.6	5.0	4.9
50 or Older	5.9	2.3	2.5
Gender			
Male	3.8	3.1	2.5
Female	5.7	5.7	4.8
Race			
White, Not Hispanic	5.1	4.5	4.3
Black, Not Hispanic	5.2	3.8	2.3
Other, Not Hispanic	9.1	4.6	2.5
Hispanic	0.5	4.8	2.6
Region			
Northeast	3.8	3.1	2.5
North Central	5.7	5.7	4.8
South	3.8	3.1	2.5
West	5.7	5.7	4.8
County Type			
Large Metro	3.8	3.1	2.5
Small Metro	5.7	5.7	4.8
Nonmetro	3.8	3.1	2.5
Received Mental Health Treatment			
Yes	28.0	18.5	17.9
No	1.1	2.3	1.5
Employment Level			
Employed Full Time	2.9	3.9	3.1
Employed Part Time	3.8	4.3	3.4
Unemployed	3.8	7.3	5.7
Others	8.2	5.1	4.7
Education			
Less Than High School	13.0	5.4	3.5
High School Graduate	1.6	4.0	3.5
Some College	8.2	6.0	4.9
College Graduate	2.7	3.0	3.0
Poverty			
<100% Threshold	20.3	7.7	5.6
100%-199% Threshold	4.9	6.3	5.1
≥200% Threshold	1.7	3.4	3.0
Health Insurance			
Yes	4.0	3.9	3.4
No	12.3	7.6	5.2

SMI =serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTES: In 2008, a split-sample design assigned adults aged 18 or older randomly to one of two impairment scales, the World Health Organization Disability Assessment Schedule (WHODAS) or the Sheehan Disability Scale (SDS). Estimates here are based on the SDS half-sample (2008B). See Table 3.2 for notes on domain definitions.

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

Table A.5 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2008 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Models
Overall	4.7	4.4	3.7
Age			
18-25	4.5	7.5	3.8
26-34	5.3	5.8	4.7
35-49	3.9	5.0	4.8
50 or Older	5.0	2.3	2.5
Gender			
Male	3.7	3.0	2.6
Female	5.5	5.7	4.7
Race			
White, Not Hispanic	5.6	4.7	4.2
Black, Not Hispanic	3.7	3.5	2.5
Other, Not Hispanic	4.8	3.7	2.7
Hispanic	0.5	4.0	2.6
Region			
Northeast	5.0	4.4	3.9
North Central	2.5	4.9	4.0
South	4.7	4.1	3.4
West	6.7	4.5	3.8
County Type			
Large Metro	2.9	4.3	3.6
Small Metro	6.6	4.5	4.0
Nonmetro	5.1	4.6	3.6
Received Mental Health Treatment			
Yes	22.2	19.1	18.1
No	1.3	2.1	1.5
Employment Level			
Employed Full Time	2.3	3.5	2.9
Employed Part Time	4.9	4.8	3.7
Unemployed	10.9	8.1	6.7
Others	9.0	5.3	4.8
Education			
Less Than High School	11.9	4.9	3.3
High School Graduate	3.7	4.4	3.9
Some College	5.6	5.5	4.6
College Graduate	2.9	3.0	2.9
Poverty			
<100% Threshold	14.3	7.4	5.9
100%-199% Threshold	8.3	6.1	4.9
≥200% Threshold	2.4	3.4	3.0
Health Insurance			
Yes	4.0	4.0	3.5
No	9.2	6.7	4.9

SMI =serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTES: In 2008, a split-sample design assigned adults aged 18 or older randomly to one of two impairment scales, the World Health Organization Disability Assessment Schedule (WHODAS) or the Sheehan Disability Scale (SDS). Estimates here are computed using both half-samples.

See Table 3.2 for notes on domain definitions.

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

Table A.6 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2009 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	5.6	4.8	3.7
Age			
18-25	3.6	7.4	3.3
26-34	3.3	6.4	5.0
35-49	13.9	5.7	4.8
50 or Older	2.5	2.8	2.5
Gender			
Male	2.7	3.2	2.6
Female	8.4	6.4	4.7
Race			
White, Not Hispanic	6.9	5.3	4.1
Black, Not Hispanic	3.0	3.7	2.7
Other, Not Hispanic	7.4	3.8	3.0
Hispanic	0.9	4.0	2.9
Region			
Northeast	0.8	4.3	3.5
North Central	3.2	5.4	4.3
South	10.1	4.9	3.4
West	6.0	4.7	3.7
County Type			
Large Metro	5.7	4.4	3.5
Small Metro	6.6	5.2	3.7
Nonmetro	4.1	5.7	4.3
Received Mental Health Treatment			
Yes	35.9	21.8	18.4
No	1.3	2.2	1.4
Employment Level			
Employed Full Time	4.3	3.6	2.8
Employed Part Time	3.1	5.6	3.5
Unemployed	8.9	7.1	5.4
Others	10.7	6.1	4.9
Education			
Less Than High School	6.6	5.7	3.9
High School Graduate	8.6	5.0	3.8
Some College	6.0	5.9	4.2
College Graduate	2.6	3.3	3.0
Poverty			
<100% Threshold	27.8	9.1	6.5
100%-199% Threshold	3.9	6.1	4.5
≥200% Threshold	3.2	3.7	2.9
Health Insurance			
Yes	5.2	4.5	3.5
No	9.0	6.5	4.8

SMI =serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.

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

Table A.7 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2010 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Model	Standard Cut Point Estimate Using 2012 Model
Overall	3.2	5.0	4.1
Age			
18-25	2.2	7.7	3.9
26-34	4.6	6.5	5.1
35-49	3.8	5.5	5.3
50 or Older	2.7	3.2	3.0
Gender			
Male	1.3	3.4	3.0
Female	5.0	6.6	5.1
Race			
White, Not Hispanic	3.5	5.3	4.3
Black, Not Hispanic	3.2	4.4	3.9
Other, Not Hispanic	0.6	4.2	3.6
Hispanic	3.2	4.6	3.2
Region			
Northeast	2.9	5.0	3.9
North Central	4.3	5.3	4.6
South	2.6	4.8	3.7
West	3.2	5.1	4.3
County Type			
Large Metro	4.2	4.7	3.8
Small Metro	1.4	5.2	4.2
Nonmetro	2.7	5.8	4.5
Received Mental Health Treatment			
Yes	20.1	22.2	20.0
No	0.3	2.3	1.5
Employment Level			
Employed Full Time	1.3	3.5	3.1
Employed Part Time	3.4	5.7	4.1
Unemployed	9.8	7.8	6.3
Others	4.7	6.6	5.3
Education			
Less Than High School	0.4	4.6	3.4
High School Graduate	1.7	5.7	4.7
Some College	6.2	5.9	4.6
College Graduate	3.5	3.6	3.3
Poverty			
<100% Threshold	4.6	9.1	7.0
100%-199% Threshold	6.0	6.0	4.7
≥200% Threshold	2.3	3.9	3.3
Health Insurance			
Yes	3.0	4.8	3.9
No	4.3	6.1	4.9

SMI =serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.

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

Table A.8 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2011 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Model	Standard Cut Point Estimate Using 2012 Model
Overall	3.7	5.0	3.9
Age			
18-25	2.9	7.6	3.8
26-34	3.3	6.7	5.0
35-49	6.4	5.6	5.1
50 or Older	2.5	3.0	2.8
Gender			
Male	3.2	3.4	2.9
Female	4.1	6.4	4.8
Race			
White, Not Hispanic	3.7	5.5	4.4
Black, Not Hispanic	2.6	3.5	2.8
Other, Not Hispanic	4.9	5.1	3.1
Hispanic	3.7	3.7	2.5
Region			
Northeast	2.7	4.1	3.4
North Central	3.1	5.1	4.2
South	3.3	4.7	3.6
West	5.8	5.9	4.4
County Type			
Large Metro	4.0	4.7	3.4
Small Metro	3.6	5.4	4.5
Nonmetro	2.9	5.1	4.4
Received Mental Health Treatment			
Yes	17.2	21.7	18.5
No	1.8	2.3	1.6
Employment Level			
Employed Full Time	2.4	3.7	2.7
Employed Part Time	4.5	5.8	4.6
Unemployed	2.9	7.6	5.4
Others	5.4	6.1	5.2
Education			
Less Than High School	7.1	5.7	4.7
High School Graduate	3.9	4.8	3.4
Some College	2.9	6.2	4.7
College Graduate	2.8	3.6	3.2
Poverty			
<100% Threshold	8.0	9.6	7.6
100%-199% Threshold	5.1	6.0	4.6
≥200% Threshold	2.4	3.6	2.8
Health Insurance			
Yes	3.2	4.6	3.6
No	6.1	6.6	5.2

SMI =serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.

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

Table A.9 Comparing SMI Prevalence Estimates Using the 2008 and 2012 Models: 2012 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	4.1	5.3	4.1
Age			
18-25	4.9	8.1	4.1
26-34	4.8	7.0	5.2
35-49	5.5	5.7	5.2
50 or Older	2.6	3.4	3.0
Gender			
Male	3.2	4.0	3.2
Female	4.9	6.5	4.9
Race			
White, Not Hispanic	4.7	5.4	4.2
Black, Not Hispanic	4.4	4.6	3.4
Other, Not Hispanic	5.2	4.8	2.9
Hispanic	0.5	5.7	4.4
Region			
Northeast	3.1	4.6	3.7
North Central	5.3	5.0	3.9
South	3.9	5.5	4.1
West	3.9	5.7	4.5
County Type			
Large Metro	3.2	5.0	3.8
Small Metro	4.9	5.8	4.5
Nonmetro	5.4	5.3	4.1
Received Mental Health Treatment			
Yes	18.5	21.6	17.6
No	1.7	2.5	1.8
Employment Level			
Employed Full Time	2.6	3.7	2.7
Employed Part Time	4.1	5.7	3.9
Unemployed	4.7	8.7	7.8
Others	6.3	7.1	5.6
Education			
Less Than High School	5.5	6.5	4.8
High School Graduate	4.0	5.6	4.4
Some College	5.0	6.0	4.4
College Graduate	2.7	3.7	3.1
Poverty			
<100% Threshold	10.4	9.8	7.2
100%-199% Threshold	4.7	6.4	5.2
≥200% Threshold	2.6	3.8	3.0
Health Insurance			
Yes	3.9	4.8	3.7
No	5.0	7.7	6.0

SMI =serious mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.

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

Table A.10 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008A NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	20.5	19.6	17.7
Age			
18-25	24.4	31.2	18.8
26-34	25.4	22.5	20.8
35-49	27.1	20.1	20.1
50 or Older	12.6	13.9	14.4
Gender			
Male	18.0	15.2	13.4
Female	22.7	23.6	21.6
Race			
White, Not Hispanic	26.3	20.0	18.4
Black, Not Hispanic	10.3	18.8	15.8
Other, Not Hispanic	3.9	20.2	16.7
Hispanic	7.0	17.9	15.9
Region			
Northeast	19.1	19.5	16.7
North Central	25.6	20.2	18.6
South	21.4	19.3	17.5
West	15.4	19.5	17.9
County Type			
Large Metro	25.2	18.5	16.9
Small Metro	14.2	20.2	18.1
Nonmetro	21.2	21.9	19.4
Received Mental Health Treatment			
Yes	53.1	54.1	53.9
No	13.1	14.2	12.1
Employment Level			
Employed Full Time	14.4	16.4	15.0
Employed Part Time	36.3	25.4	20.9
Unemployed	54.6	29.6	23.7
Others	26.7	21.5	20.6
Education			
Less Than High School	35.1	21.8	19.1
High School Graduate	21.2	19.4	17.8
Some College	16.4	21.1	18.5
College Graduate	20.3	17.2	16.0
Poverty			
<100% Threshold	36.9	29.4	25.0
100%-199% Threshold	27.7	25.2	22.4
≥200% Threshold	17.1	16.4	15.3
Health Insurance			
Yes	21.4	18.5	17.0
No	16.3	25.4	21.4

AMI =any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTES: In 2008, a split-sample design assigned adults aged 18 or older randomly to one of two impairment scales, the World Health Organization Disability Assessment Schedule (WHODAS) or the Sheehan Disability Scale (SDS). Estimates here are based on the WHODAS half-sample (2008A). See Table 3.2 for notes on domain definitions.

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

Table A.11 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008B NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	19.1	16.9	17.8
Age			
18-25	19.9	26.8	18.4
26-34	28.0	20.5	23.1
35-49	21.5	18.2	20.2
50 or Older	13.4	11.0	13.8
Gender			
Male	15.2	13.3	13.8
Female	22.8	20.2	21.5
Race			
White, Not Hispanic	18.2	16.9	18.1
Black, Not Hispanic	24.8	17.1	17.6
Other, Not Hispanic	26.1	17.1	16.9
Hispanic	15.6	16.4	16.7
Region			
Northeast	21.8	19.0	21.2
North Central	10.6	17.3	18.0
South	22.5	15.6	16.5
West	20.3	16.8	16.9
County Type			
Large Metro	19.9	16.8	17.6
Small Metro	20.2	17.6	18.6
Nonmetro	15.8	15.9	16.9
Received Mental Health Treatment			
Yes	57.2	48.8	53.9
No	13.2	11.9	12.2
Employment Level			
Employed Full Time	19.8	15.0	16.4
Employed Part Time	24.4	18.0	17.0
Unemployed	26.9	25.8	25.5
Others	15.3	18.6	19.7
Education			
Less Than High School	29.4	21.1	21.6
High School Graduate	11.3	15.8	16.3
Some College	28.8	19.7	20.1
College Graduate	15.8	13.1	15.1
Poverty			
<100% Threshold	36.1	25.0	25.7
100%-199% Threshold	23.1	20.9	21.1
≥200% Threshold	14.9	14.4	15.6
Health Insurance			
Yes	17.9	15.6	16.6
No	31.9	24.2	24.4

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTES: In 2008, a split-sample design assigned adults aged 18 or older randomly to one of two impairment scales, the World Health Organization Disability Assessment Schedule (WHODAS) or the Sheehan Disability Scale (SDS). Estimates here are based on the SDS half-sample (2008B). See Table 3.2 for notes on domain definitions.

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

Table A.12 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2008 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	19.8	18.2	17.7
Age			
18-25	22.2	28.9	18.5
26-34	26.6	21.6	21.9
35-49	24.3	19.1	20.1
50 or Older	13.0	12.4	14.1
Gender			
Male	16.6	14.2	13.6
Female	22.7	21.9	21.5
Race			
White, Not Hispanic	22.3	18.4	18.3
Black, Not Hispanic	17.6	18.0	16.7
Other, Not Hispanic	15.0	18.7	16.8
Hispanic	11.3	17.0	16.0
Region			
Northeast	20.6	19.2	18.9
North Central	17.6	18.7	18.4
South	21.9	17.4	17.0
West	17.6	18.1	17.3
County Type			
Large Metro	22.6	17.6	17.2
Small Metro	17.2	19.0	18.4
Nonmetro	18.5	18.7	18.0
Received Mental Health Treatment			
Yes	54.8	51.4	53.8
No	13.2	13.0	12.1
Employment Level			
Employed Full Time	16.8	15.7	15.7
Employed Part Time	30.0	21.6	18.8
Unemployed	38.6	27.5	24.7
Others	19.7	20.0	20.0
Education			
Less Than High School	32.2	21.3	20.3
High School Graduate	15.7	17.5	17.0
Some College	22.5	20.3	19.3
College Graduate	18.3	15.3	15.6
Poverty			
<100% Threshold	36.5	26.9	25.2
100%-199% Threshold	25.0	22.9	21.6
≥200% Threshold	16.1	15.4	15.5
Health Insurance			
Yes	19.5	17.0	16.8
No	21.4	24.7	22.7

AMI =any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTES: In 2008, a split-sample design assigned adults aged 18 or older randomly to one of two impairment scales, the World Health Organization Disability Assessment Schedule (WHODAS) or the Sheehan Disability Scale (SDS). Estimates here are computed using both half samples. See Table 3.2 for notes on domain definitions.

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

Table A.13 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2009 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	20.3	20.0	18.1
Age			
18-25	21.5	30.2	18.0
26-34	18.3	24.5	22.8
35-49	26.5	21.1	20.9
50 or Older	17.1	13.8	14.5
Gender			
Male	11.8	15.7	14.2
Female	28.2	24.0	21.8
Race			
White, Not Hispanic	23.0	20.8	19.2
Black, Not Hispanic	13.6	18.0	15.6
Other, Not Hispanic	20.9	19.1	16.8
Hispanic	12.0	17.8	15.7
Region			
Northeast	5.3	19.7	18.0
North Central	13.6	20.3	18.4
South	24.1	19.6	17.6
West	35.3	20.5	18.8
County Type			
Large Metro	20.6	19.6	17.7
Small Metro	25.9	20.4	18.7
Nonmetro	11.2	20.5	18.4
Received Mental Health Treatment			
Yes	77.1	56.6	54.4
No	12.2	14.3	12.5
Employment Level			
Employed Full Time	18.6	17.2	16.1
Employed Part Time	17.3	23.6	19.0
Unemployed	53.1	27.6	23.2
Others	19.2	21.3	20.2
Education			
Less Than High School	17.9	21.9	18.9
High School Graduate	28.8	19.5	17.6
Some College	21.9	22.6	20.4
College Graduate	13.2	17.1	16.3
Poverty			
<100% Threshold	52.4	30.0	25.1
100%-199% Threshold	15.5	22.6	20.2
≥200% Threshold	17.5	17.3	16.3
Health Insurance			
Yes	19.2	19.0	17.7
No	29.6	25.0	20.5

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.

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

Table A.14 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2010 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	17.7	20.1	18.1
Age			
18-25	19.8	30.1	18.1
26-34	14.5	24.7	21.6
35-49	25.0	20.7	20.6
50 or Older	13.2	14.4	15.1
Gender			
Male	18.1	16.8	14.8
Female	17.4	23.1	21.1
Race			
White, Not Hispanic	18.6	20.7	19.0
Black, Not Hispanic	9.4	19.8	16.9
Other, Not Hispanic	34.9	18.2	15.9
Hispanic	12.6	18.3	15.2
Region			
Northeast	19.8	20.4	18.7
North Central	17.2	20.7	18.7
South	12.7	19.7	17.4
West	22.3	19.8	17.9
County Type			
Large Metro	17.0	19.5	17.5
Small Metro	23.0	20.8	18.8
Nonmetro	9.3	20.7	18.7
Received Mental Health Treatment			
Yes	54.8	57.3	55.6
No	11.3	14.1	12.1
Employment Level			
Employed Full Time	19.7	16.8	15.3
Employed Part Time	13.9	22.9	18.7
Unemployed	19.8	27.9	22.7
Others	14.8	22.6	21.4
Education			
Less Than High School	8.5	22.2	18.5
High School Graduate	16.8	20.3	18.1
Some College	19.9	21.9	19.4
College Graduate	21.1	17.1	16.7
Poverty			
<100% Threshold	22.1	29.6	24.7
100%-199% Threshold	26.1	23.5	20.5
≥200% Threshold	15.1	17.1	16.0
Health Insurance			
Yes	15.1	19.2	17.6
No	34.5	24.5	20.3

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.
Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health.

Table A.15 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2011 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	17.1	19.6	17.8
Age			
18-25	19.3	29.8	18.5
26-34	20.4	23.1	20.9
35-49	20.2	20.4	20.0
50 or Older	13.1	14.3	15.0
Gender			
Male	13.2	15.9	14.2
Female	20.6	23.0	21.1
Race			
White, Not Hispanic	17.1	20.5	19.0
Black, Not Hispanic	15.3	18.8	16.3
Other, Not Hispanic	12.9	20.2	17.3
Hispanic	20.1	15.9	13.5
Region			
Northeast	13.5	18.4	16.3
North Central	17.6	20.1	18.0
South	15.8	19.4	17.9
West	22.1	20.4	18.6
County Type			
Large Metro	18.7	18.5	16.8
Small Metro	15.5	20.8	18.8
Nonmetro	15.1	20.9	19.1
Received Mental Health Treatment			
Yes	49.7	54.9	53.2
No	12.6	14.0	12.2
Employment Level			
Employed Full Time	14.7	16.1	14.5
Employed Part Time	19.8	23.2	20.1
Unemployed	13.8	28.0	23.4
Others	20.2	22.0	20.9
Education			
Less than High School	22.0	22.5	19.9
High School Graduate	17.4	18.9	17.0
Some College	14.0	21.8	19.3
College Graduate	17.6	16.8	16.3
Poverty			
<100% Threshold	23.3	29.6	25.8
100%-199% Threshold	21.6	23.1	20.5
≥200% Threshold	14.6	16.2	15.2
Health Insurance			
Yes	16.3	18.7	17.1
No	21.0	24.3	21.4

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.

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

Table A.16 Comparing AMI Prevalence Estimates Using the 2008 and 2012 Models: 2012 NSDUH by Domain

Domain	Direct Estimate from MHSS Clinical Subsample	Standard Cut Point Estimate Using 2008 Models	Standard Cut Point Estimate Using 2012 Model
Overall	18.1	20.5	18.6
Age			
18-25	22.6	31.0	19.6
26-34	18.3	25.5	22.9
35-49	16.1	20.3	20.2
50 or Older	17.6	15.3	15.8
Gender			
Male	13.6	16.9	14.9
Female	22.3	23.9	22.0
Race			
White, Not Hispanic	17.1	20.8	19.3
Black, Not Hispanic	18.5	21.6	18.6
Other, Not Hispanic	13.1	18.9	17.0
Hispanic	24.4	19.3	16.3
Region			
Northeast	25.9	20.0	18.0
North Central	14.9	20.1	18.2
South	17.0	20.5	18.7
West	16.0	21.5	19.3
County Type			
Large Metro	19.8	20.0	18.0
Small Metro	15.0	21.5	19.6
Nonmetro	17.8	20.3	18.7
Received Mental Health Treatment			
Yes	52.6	54.1	52.4
No	12.5	14.8	12.9
Employment Level			
Employed Full Time	12.2	16.7	15.2
Employed Part Time	18.6	23.7	19.8
Unemployed	24.4	30.6	25.5
Others	26.5	23.5	22.3
Education			
Less than High School	35.3	24.4	21.9
High School Graduate	16.5	20.8	18.7
Some College	14.7	22.2	19.7
College Graduate	14.1	16.8	15.9
Poverty			
<100% Threshold	23.2	30.8	26.8
100%-199% Threshold	27.0	23.9	21.8
≥200% Threshold	14.4	16.9	15.6
Health Insurance			
Yes	17.1	19.7	17.9
No	22.3	24.9	22.3

AMI = any mental illness; NSDUH = National Survey on Drug Use and Health; MHSS = Mental Health Surveillance Study.

NOTE: See Table 3.2 for notes on domain definitions.

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