

NATIONAL SURVEY ON DRUG USE AND HEALTH: ALTERNATIVE STATISTICAL MODELS TO PREDICT MENTAL ILLNESS

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
Center for Behavioral Health Statistics and Quality
Rockville, Maryland 20857

September 2015

This page intentionally left blank

NATIONAL SURVEY ON DRUG USE AND HEALTH: ALTERNATIVE STATISTICAL MODELS TO PREDICT MENTAL ILLNESS

Contract No. HHSS283201000003C
RTI Project No. 0212800.001.112.008.009.006

RTI Authors:

Jeremy Aldworth
Kristen Gulledge
Lauren Warren

RTI Project Director:

David Hunter

SAMHSA Project Officer:

Peter Tice

SAMHSA Authors:

Sarra Hedden
Jonaki Bose

For questions about this report, please e-mail 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). *National Survey on Drug Use and Health: Alternative statistical
models to predict mental illness*. Rockville, MD: Substance Abuse and Mental
Health Services Administration.

Acknowledgments

This methodological report was prepared for the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality (CBHSQ), by RTI International (a registered trademark and a trade name of Research Triangle Institute). Significant other contributors to this report at SAMHSA included Joe Gfroerer. At RTI, the report was edited by Richard S. Straw and formatted by Debbie Bond.

Table of Contents

Chapter	Page
1. Introduction.....	1
2. Estimating SMI and AMI from the 2012 SMI Prediction Model.....	3
3. Parsimonious Alternative Models for Estimating SMI and AMI.....	5
3.1 Description of Parsimonious Models.....	5
3.2 ROC Analysis of Parsimonious Models.....	8
4. SMI and AMI National Estimates and Domain-Level Bias among NSDUH Models.....	13
4.1 Domain-Level Bias of Estimates among Parsimonious Models.....	13
4.2 Comparisons of SMI and AMI National Estimates among NSDUH Models.....	22
5. Analyzing SMI or AMI in Conjunction with Model Predictor Variables.....	29
6. Comparisons of SMI and AMI National Estimates between NSDUH and Other Data Sources.....	33
6.1 Short Descriptions of NSDUH and Other Data Sources.....	33
6.1.1 National Survey on Drug Use and Health (NSDUH).....	33
6.1.2 Behavioral Risk Factor Surveillance System (BRFSS).....	34
6.1.3 National Health Interview Survey (NHIS).....	34
6.1.4 Medical Expenditure Panel Survey (MEPS).....	34
6.2 Comparisons of SMI and AMI Estimates between NSDUH and Other Data Sources.....	35
7. Summary and Conclusions.....	43
References.....	47

This page intentionally left blank

List of Tables

Table	Page
3.1 Predictor Variables of Models of Varying Parsimony.....	5
3.2 Beta Coefficients and Fit Statistics of Models That Include K6 (Past Year)	6
3.3 Beta Coefficients and Fit Statistics of Models That Include K6 (Past Month)	7
3.4 ROC Statistics of SMI Cut Point Estimates Based on Parsimonious Models	9
3.5 ROC Statistics of AMI Cut Point Estimates Based on Parsimonious Models	10
4.1 Domain-Level Bias of SMI Cut Point Estimates Based on Models That Include K6 (Past Year) Term: 2008A-2012	14
4.2 Domain-Level Bias of AMI Cut Point Estimates Based on Models That Include K6 (Past Year) Term: 2008A-2012	16
4.3 Domain-Level Bias of SMI Cut Point Estimates Based on Models That Include K6 (Past Month) Term: 2008A-2012.....	18
4.4 Domain-Level Bias of AMI Cut Point Estimates Based on Models That Include K6 (Past Month) Term: 2008A-2012.....	20
4.5 NSDUH Estimates of SMI Based on Models of Varying Parsimony, by Demographic Domain: 2008A-2012.....	24
4.6 NSDUH Estimates of AMI Based on Models of Varying Parsimony, by Demographic Domain: 2008A-2012.....	26
5.1 Comparing SMI and AMI Prevalence Estimates for MDE and Suicidal Thought Domains Using Models Excluding or Including Past Year MDE or Suicidal Thoughts as Predictor Variables	30
5.2 Comparing Prevalence Estimates of Suicidal Thoughts, Plans, and Attempts in the Past Year among Adults with SMI Using Models Excluding or Including Suicidal Thoughts as a Predictor Variable.....	32
6.1 Predictor Variables Included in NSDUH and Other Studies	33
6.2 Comparison of Estimates of SMI between Selected NSDUH and BRFSS Models: 2012.....	37
6.3 Comparison of Estimates of SMI between Selected NSDUH and NHIS Models: 2012.....	38
6.4 Comparison of Estimates of SMI between Selected NSDUH and MEPS Models: 2010.....	39
6.5 Comparison of Estimates of AMI between Selected NSDUH and BRFSS Models: 2012.....	40
6.6 Comparison of Estimates of AMI between Selected NSDUH and NHIS Models: 2012.....	41
6.7 Comparison of Estimates of AMI between Selected NSDUH and MEPS Models: 2010.....	42
7.1 Strengths and Weakness of Models	45

This page intentionally left blank

1. Introduction

The Mental Health Surveillance Study (MHSS) clinical follow-up was conducted as part of the National Survey on Drug Use and Health (NSDUH) from 2008 to 2012 for the primary purpose of developing models that estimate the prevalence of serious mental illness (SMI) and any mental illness (AMI) in the adult (18 or older) U.S. population (Center for Behavioral Health Statistics and Quality [CBHSQ], 2014). In each of these years, the MHSS clinical study consisted of a subsample of eligible adults selected from the main NSDUH study for follow-up clinical interviews. A prediction model was developed that included SMI status (collected in the clinical interview) as the dependent variable. The predictor variables were variables such as psychological distress and impairment measures that were collected in the main NSDUH questionnaire for adults. The resulting prediction model developed from the MHSS clinical data was then applied to all adult NSDUH respondents to obtain the predicted probability of SMI for each respondent. Next, a cut point was determined so that if the predicted probability of SMI for a respondent met or exceeded the cut point, then he or she was *predicted* to be SMI positive; otherwise, he or she was predicted to be SMI negative. That is, a dichotomy of *SMI status* was computed. Receiver operating characteristic (ROC) analysis was applied to determine the cut point that resulted in the weighted number of false-positive and false-negative counts being (at least approximately) equal, thus ensuring unbiased estimates.

In 2008, when only 750 clinical interview cases were available, an SMI prediction model was constructed (CBHSQ, 2009) that contained only two terms: (1) a past year Kessler-6 (K6) score (Kessler et al., 2003), and (2) an abbreviated eight-item version of the World Health Organization Disability Assessment Schedule (WHODAS) (Novak, Colpe, Barker, & Gfroerer, 2010; Rehm et al., 1999).

By 2012, the combined 2008 to 2012 WHODAS clinical sample included approximately 5,000 respondents. Based on this much larger sample, the SMI prediction model was revised and improved. As a result of the revision, the 2012 SMI prediction model added three terms related to age, past year major depression episode (MDE), and suicidal thoughts to the two terms already in the 2008 model. Analyses indicated that some of these added terms reduced bias at the domain (i.e., subpopulation) level of various demographic and geographic variables, and others reduced the overall error rate (i.e., the sum of false-positive and false-negative rates).

However, a couple of issues are related to the more complex 2012 prediction model:

- Researchers should not analyze SMI prevalence within domains of variables that are themselves predictor variables in the model used to predict SMI, or closely related to those predictor variables, because in these circumstances SMI estimates tend to be biased. For example, researchers should not cross SMI by past year MDE or suicidal thoughts if they used the 2012 prediction model to predict SMI because in this case SMI tends to be overestimated among individuals with past year MDE or suicidal thoughts and underestimated among those without past year MDE or suicidal thoughts. Instead, researchers should use a model that does not include past year MDE or suicidal thoughts as a predictor variable.

- There is an interest in using this model to estimate SMI based on other sources of data (e.g., for specific subpopulations or populations not covered by NSDUH). However, other data sources that do not collect all of the predictor variables would not be able to use this model to predict SMI. For example, although the Behavioral Risk Factor Surveillance System (BRFSS), the National Health Interview Survey (NHIS), and the Medical Expenditure Panel Survey (MEPS) all collected past month K6 data (compared with NSDUH that collects both past month and past year), none of these sources collected information resembling the WHODAS scale, past year MDE, or past year suicidal thoughts.

Therefore, the objective of this methods study was to use the MHSS clinical data to characterize more parsimonious SMI prediction models in terms of bias, classification error rate, and how they compare with the 2012 model to ascertain their usefulness in situations where (1) one of the predictor variables in the SMI prediction model happens to be a key analytic variable with respect to SMI, and (2) not all the information required to construct the 2012 model is available.

It cannot be assumed, however, that models derived from NSDUH data are equally valid for use based on other sources of data. Therefore, comparisons of SMI and AMI estimates between those derived from NSDUH and those derived from BRFSS, NHIS, and MEPS data sources are also included using a few of these parsimonious models. Observed differences in the comparisons could be due to a multitude of reasons, including the different methodological characteristics (e.g., variable collection and survey methods) used for each data source. Therefore, a brief description is also provided of the different methodological characteristics of NSDUH and the other three data sources.

This methodological document is organized as follows. As a reference, the 2012 SMI prediction model and method for determining estimates of SMI and AMI is briefly discussed in Chapter 2. Parsimonious alternative models for determining estimates of SMI and AMI, including caveats about which models should not be used, are described in Chapter 3. Comparisons of national estimates of SMI and AMI and domain-level bias among the 2012 and alternative models are detailed in Chapter 4. An analysis of SMI or AMI in conjunction with model predictor variables is discussed in Chapter 5. Comparisons in SMI and AMI national estimates between those derived from NSDUH and those derived from BRFSS, NHIS, and MEPS data sources are described in Chapter 6. Concluding remarks are given in Chapter 7.

2. Estimating SMI and AMI from the 2012 SMI Prediction Model

The 2012 serious mental illness (SMI) prediction model is used to provide national estimates of SMI and any mental illness (AMI) based on data from the 2008 to 2012 National Surveys on Drug Use and Health (NSDUHs) for adults aged 18 or older who were assigned to the World Health Organization Disability Assessment Schedule (WHODAS) questions¹ (Center for Behavioral Health Statistics and Quality [CBHSQ], 2014). Specifically, let π represent the probability that an adult has SMI. Then the 2012 model can be expressed as follows:

$$\text{logit}(\hat{\pi}) \equiv \log[\hat{\pi} / (1 - \hat{\pi})] = -5.9726644 + 0.0873416 * X_k + 0.3385193 * X_w + 1.9552664 * X_s + 1.1267330 * X_m + 0.1059137 * X_a$$

where $\hat{\pi}$ is the estimated probability that an adult had SMI, and the covariates X_k , X_w , X_s , X_m , and X_a are defined as follows:

X_k = *Alternative Past Year Kessler-6 (K6) Score*: Past year K6 score of less than 8 recoded as 0; past year K6 score of 8 to 24 recoded as 1 to 17.

X_w = *Alternative WHODAS Score*: WHODAS item score of less than 2 recoded as 0; WHODAS item score of 2 to 3 recoded as 1, then summed for a score ranging from 0 to 8.

X_s = *Serious Thoughts of Suicide in the Past Year*: Coded as 1 if "yes"; coded as 0 otherwise.

X_m = *Past Year Major Depressive Episode (MDE) on NSDUH*: Coded as 1 if the criteria for past year MDE were met;² coded as 0 otherwise.

X_a = *Adjusted Age (i.e., AGE1830)*: Coded as age minus 18 if aged 18 to 30; coded as 12 otherwise.

The 2012 formula for the predicted probability of SMI (SMIPP_U) can then be expressed using the model parameter estimates above as follows:

$$\text{SMIPP_U} = 1 / (1 + \exp[-(-5.9726644 + 0.0873416 * X_k + 0.3385193 * X_w + 1.9552664 * X_s + 1.1267330 * X_m + 0.1059137 * X_a)]).$$

If SMIPP_U was greater than or equal to 0.2605735292 (the SMI cut point), then the respondent was predicted as having past year SMI (i.e.; SMIYR_U = 1); otherwise, the respondent was predicted as not having past year SMI (SMIYR_U = 0). If SMIPP_U was greater than or equal to 0.0192519810 (the AMI cut point), then the respondent was predicted as having

¹ Half of the adult respondents in the 2008 NSDUH were assigned to the WHODAS questions, and the other half were assigned to questions associated with the Sheehan Disability Scale (SDS) (Leon, Olfson, Portera, Farber, & Sheehan, 1997). A separate model was developed for those assigned to the SDS.

² See Section B.4.5 of CBHSQ (2015) for a detailed discussion of the past year MDE criteria.

past year AMI (AMIYR_U = 1); otherwise, the respondent was predicted as not having past year AMI (AMIYR_U = 0).

3. Parsimonious Alternative Models for Estimating SMI and AMI

3.1 Description of Parsimonious Models

A set of models with varying degrees of parsimony was selected for further investigation to see how they would perform in terms of bias (overall and at the domain level), error rate (sum of false-positive and false-negative rates), and how they would compare against the 2012 model. One set of six models includes the past year version of the Kessler-6 (K6) scale, and a second set of six models is identical except that the past month version of the K6 scale is used instead. The reason for the second set of models is that the past month version of the K6 scale is collected in surveys such as the Behavioral Risk Factor Surveillance System (BRFSS), the National Health Interview Survey (NHIS), and the Medical Expenditure Panel Survey (MEPS). Model Y6 (shown in [Table 3.1](#)) represents the current model used to predict serious mental illness (SMI) on the National Survey on Drug Use and Health (NSDUH). The predictor variables of all 12 models that were examined for this study are listed in [Table 3.1](#).

Table 3.1 Predictor Variables of Models of Varying Parsimony

Model	Predictor Variables
Y1	K6 (Past Year)
Y2	K6 (Past Year) + Age
Y3	K6 (Past Year) + WHODAS
Y4	K6 (Past Year) + WHODAS + Age
Y5	K6 (Past Year) + WHODAS + Age + Past Year MDE
Y6	K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts
M1	K6 (Past Month)
M2	K6 (Past Month) + Age
M3	K6 (Past Month) + WHODAS
M4	K6 (Past Month) + WHODAS + Age
M5	K6 (Past Month) + WHODAS + Age + Past Year MDE
M6	K6 (Past Month) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

K6 = Kessler 6; MDE = major depressive episode; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: The following predictor variables are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Beta coefficients and their test statistics for the first six models listed in [Table 3.1](#) (i.e., Y1 to Y6) are shown in [Table 3.2](#), and similar statistics for the final six models listed in [Table 3.1](#) (i.e., M1 to M6) are shown in [Table 3.3](#). Also included in [Tables 3.2](#) and [3.3](#) are the cut points for SMI and any mental illness (AMI) prediction for each of the models.

Table 3.2 Beta Coefficients and Fit Statistics of Models That Include K6 (Past Year)

Model	Predictor Variable	DF	Beta	Beta SE	t Value	p Value	SMI Cut Point	AMI Cut Point
Y1	Intercept		-4.2546253	0.132	-32.23	0.0000	0.2438354935	0.0322038289
	K6 (Past Year)	1	0.2838964	0.015	19.57	0.0000		
Y2	Intercept		-5.1976933	0.299	-17.38	0.0000	0.2293531541	0.0271091587
	K6 (Past Year)	1	0.2933824	0.015	20.05	0.0000		
	Age	1	0.0858774	0.024	3.56	0.0008		
Y3	Intercept		-4.7115050	0.147	-32.13	0.0000	0.2682495177	0.0282057278
	K6 (Past Year)	1	0.1550075	0.019	8.11	0.0000		
	WHODAS	1	0.3968483	0.033	12.01	0.0000		
Y4	Intercept		-5.4074873	0.299	-18.10	0.0000	0.2733798039	0.0253955684
	K6 (Past Year)	1	0.1638237	0.019	8.54	0.0000		
	WHODAS	1	0.3899276	0.032	12.16	0.0000		
	Age	1	0.0649826	0.023	2.77	0.0079		
Y5	Intercept		-5.3961874	0.309	-17.48	0.0000	0.2668853750	0.0221456122
	K6 (Past Year)	1	0.1163358	0.024	4.80	0.0000		
	WHODAS	1	0.3397227	0.032	10.51	0.0000		
	Past Year MDE	1	1.3470857	0.233	5.79	0.0000		
	Age	1	0.0619774	0.025	2.50	0.0156		
Y6	Intercept		-5.9726644	0.320	-18.66	0.0000	0.2605735292	0.0192519810
	K6 (Past Year)	1	0.0873416	0.025	3.52	0.0009		
	WHODAS	1	0.3385193	0.035	9.70	0.0000		
	Suicidal Thoughts	1	1.9552664	0.216	9.03	0.0000		
	Past Year MDE	1	1.1267330	0.220	5.13	0.0000		
	Age	1	0.1059137	0.024	4.34	0.0001		

AMI = any mental illness; DF = degrees of freedom; MDE = major depressive episode; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: The following predictor variables are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The model Y6 (in bold) describes the 2012 model.

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

Table 3.3 Beta Coefficients and Fit Statistics of Models That Include K6 (Past Month)

Model	Predictor Variable	DF	Beta	Beta SE	t Value	p Value	SMI Cut Point	AMI Cut Point
M1	Intercept		-3.8155673	0.103	-37.01	0.0000	0.1847185719	0.0298117857
	K6 (Past Month)	1	0.3329811	0.021	15.52	0.0000		
M2	Intercept		-4.5039168	0.304	-14.84	0.0000	0.1553436127	0.0237243262
	K6 (Past Month)	1	0.3411437	0.022	15.60	0.0000		
	Age	1	0.0636470	0.026	2.49	0.0160		
M3	Intercept		-4.6485793	0.138	-33.70	0.0000	0.2732661731	0.0255277417
	K6 (Past Month)	1	0.1776653	0.019	9.45	0.0000		
	WHODAS	1	0.4734530	0.027	17.78	0.0000		
M4	Intercept		-5.1599999	0.291	-17.72	0.0000	0.2660257241	0.0257593844
	K6 (Past Month)	1	0.1828271	0.019	9.48	0.0000		
	WHODAS	1	0.4716675	0.026	18.09	0.0000		
	Age	1	0.0486505	0.024	2.07	0.0437		
M5	Intercept		-5.2218088	0.302	-17.32	0.0000	0.2573724276	0.0209226021
	K6 (Past Month)	1	0.1338454	0.020	6.57	0.0000		
	WHODAS	1	0.3875149	0.029	13.56	0.0000		
	Past Year MDE	1	1.4502682	0.199	7.27	0.0000		
	Age	1	0.0498655	0.024	2.04	0.0470		
M6	Intercept		-5.8972298	0.330	-17.87	0.0000	0.2618658262	0.0189881867
	K6 (Past Month)	1	0.1105975	0.022	4.97	0.0000		
	WHODAS	1	0.3703121	0.030	12.42	0.0000		
	Suicidal Thoughts	1	1.9966104	0.200	10.01	0.0000		
	Past Year MDE	1	1.1734907	0.206	5.71	0.0000		
	Age	1	0.1009016	0.025	3.99	0.0002		

AMI = any mental illness; DF = degrees of freedom; MDE = major depressive episode; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: The following predictor variables are defined in exactly the same way as in the 2012 model (see Chapter 2): WHODAS, Age, Suicidal Thoughts, and Past Year MDE.

The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead.

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

3.2 ROC Analysis of Parsimonious Models

The 12 models described in [Table 3.1](#) were run using the 2008A to 2012 Mental Health Surveillance Study (MHSS) clinical data (i.e., the World Health Organization Disability Assessment Schedule [WHODAS] sample only) and the rescaled MHSS analysis weights for the WHODAS sample (i.e., MHFAAWGT), yielding a predicted probability of SMI for each respondent and model. The next step was to determine SMI and AMI cut points for each model so that predicted SMI status and AMI status could be determined for each respondent and model.

The cut point estimation method relies on being able to identify a cut point that equalizes the weighted false-positive and false-negative rates (or counts) to provide unbiased estimates of SMI or AMI. But the variable nature of unequal weights makes it difficult to achieve exact equality between the two rates, thereby resulting in some bias. However, for a model that yields a sufficient number of distinct, realizable, and predicted probabilities, it is typically not difficult to minimize the bias to a reasonable extent. Unfortunately, for very parsimonious models that do not yield a large number of distinct, realizable, and predicted probabilities,³ the closest the weighted false-positive and false-negative rates can be equalized can still result in a fairly substantial bias. Therefore, a model that is too parsimonious may not provide reasonably unbiased estimates.

Receiver operating characteristic (ROC) statistics of the cut point estimators associated with the 12 models for both SMI and AMI are displayed in [Tables 3.4](#) and [3.5](#), respectively. The following ROC statistics are included in the tables:

- false-positive rate, which is the proportion of all respondents who were predicted to have SMI but did not have SMI;
- false-negative rate, which is the proportion of all respondents who were predicted not to have SMI but did have SMI;
- total error rate, which is the sum of the false-positive and false-negative rates;
- sensitivity, which is the proportion of respondents with SMI who were predicted to have SMI;
- specificity, which is the proportion of respondents without SMI who were predicted not to have SMI; and
- area under ROC curve based on predicted responses dichotomized by the cut point, which is the average of sensitivity and specificity.

³ For example, if a model contains a single predictor variable with six distinct levels, then the set of distinct, realizable, and predicted probabilities would also be six. If there are too few distinct, realizable, and predicted values, then the coarseness of the gradation between those values may lead to a scenario where a cut point defined one way may result in too many false positives, but then if dropped to the next level would result in too many false negatives.

Table 3.4 ROC Statistics of SMI Cut Point Estimates Based on Parsimonious Models

Model	Sample Size	Pop Size (1,000s)	Cut Point	SCID Estimate	Cut Point Estimate	SE	Bias	False-Positive Rate	False-Negative Rate	Total Error Rate	Sensitivity	Specificity	AUC
Y1	4,912	231,890	0.243835	3.93	4.14	0.336	0.2055	2.76	2.55	5.31	0.351	0.971	0.661
Y2	4,912	231,890	0.229353	3.93	3.72	0.331	-0.2141	2.40	2.61	5.01	0.335	0.975	0.655
Y3	4,912	231,890	0.268250	3.93	3.93	0.296	0.0005	2.21	2.21	4.42	0.438	0.977	0.707
Y4	4,912	231,890	0.273380	3.93	3.98	0.294	0.0531	2.17	2.12	4.30	0.460	0.977	0.719
Y5	4,912	231,890	0.273380	3.93	3.93	0.286	-0.0045	2.03	2.03	4.06	0.483	0.979	0.731
Y6	4,912	231,890	0.260574	3.93	3.92	0.270	-0.0128	1.92	1.93	3.84	0.509	0.980	0.745
M1	4,912	231,890	0.184719	3.93	3.46	0.331	-0.4658	2.23	2.70	4.93	0.314	0.977	0.645
M2	4,912	231,890	0.155344	3.93	3.96	0.334	0.0296	2.55	2.52	5.08	0.358	0.973	0.666
M3	4,912	231,890	0.273266	3.93	3.92	0.310	-0.0122	2.33	2.34	4.68	0.403	0.976	0.690
M4	4,912	231,890	0.266026	3.93	3.95	0.308	0.0151	2.35	2.34	4.69	0.405	0.976	0.690
M5	4,912	231,890	0.266026	3.93	3.93	0.281	0.0003	1.96	1.96	3.92	0.501	0.980	0.740
M6	4,912	231,890	0.261866	3.93	3.94	0.269	0.0072	1.88	1.88	3.76	0.523	0.980	0.752

AUC = area under ROC curve; K6 = Kessler-6; MDE = major depressive episode; Pop = population; ROC = receiver operating characteristic; SCID = Structured Clinical Interview for DSM-IV; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: Bias = Difference between cut point estimate of SMI from indicated model and direct estimate of SMI as computed in the clinical sample.

Predictor variables included in the following models:

- Y1: K6 (Past Year)
- Y2: K6 (Past Year) + Age
- Y3: K6 (Past Year) + WHODAS
- Y4: K6 (Past Year) + WHODAS + Age
- Y5: K6 (Past Year) + WHODAS + Age + Past Year MDE
- Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts**
- M1: K6 (Past Month)
- M2: K6 (Past Month) + Age
- M3: K6 (Past Month) + WHODAS
- M4: K6 (Past Month) + WHODAS + Age
- M5: K6 (Past Month) + WHODAS + Age + Past Year MDE
- M6: K6 (Past Month) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Data and weights:

- Dataset = 2008A-2012 MHSS clinical data.
- Analysis weight = MHFAAWGT.

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

Table 3.5 ROC Statistics of AMI Cut Point Estimates Based on Parsimonious Models

Model	Sample Size	Pop Size (1,000s)	Cut Point	SCID Estimate	Cut Point Estimate	SE	Bias	False Positive Rate	False Negative Rate	Total Error Rate	Sensitivity	Specificity	AUC
Y1	4,912	231,890	0.032204	18.03	17.47	1.177	-0.5624	8.43	8.99	17.43	0.501	0.897	0.699
Y2	4,912	231,890	0.027109	18.03	18.29	1.173	0.2586	8.73	8.47	17.21	0.530	0.893	0.712
Y3	4,912	231,890	0.028206	18.03	18.15	1.112	0.1263	7.99	7.86	15.85	0.564	0.903	0.733
Y4	4,912	231,890	0.025396	18.03	18.03	1.112	0.0044	7.79	7.79	15.58	0.568	0.905	0.736
Y5	4,912	231,890	0.022146	18.03	18.03	1.113	0.0052	7.73	7.72	15.45	0.572	0.906	0.739
Y6	4,912	231,890	0.019252	18.03	17.96	1.112	-0.0646	7.70	7.77	15.47	0.569	0.906	0.738
M1	4,912	231,890	0.029812	18.03	15.34	1.206	-2.6902	7.48	10.17	17.65	0.436	0.909	0.672
M2	4,912	231,890	0.023724	18.03	14.52	1.199	-3.5077	6.87	10.38	17.25	0.424	0.916	0.670
M3	4,912	231,890	0.025528	18.03	17.05	1.107	-0.9761	7.37	8.35	15.72	0.537	0.910	0.724
M4	4,912	231,890	0.025759	18.03	18.98	1.126	0.9569	8.64	7.68	16.32	0.574	0.895	0.734
M5	4,912	231,890	0.020923	18.03	17.14	1.109	-0.8891	7.23	8.12	15.34	0.550	0.912	0.731
M6	4,912	231,890	0.018988	18.03	16.95	1.106	-1.0809	7.07	8.15	15.23	0.548	0.914	0.731

AMI = any mental illness; AUC = area under ROC curve; K6 = Kessler-6; MDE = major depressive episode; Pop = population; ROC = receiver operating characteristic; SCID = Structured Clinical Interview for DSM-IV; SE = standard error; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: Bias = Difference between cut point estimate of AMI from indicated model and direct estimate of AMI as computed in the clinical sample.

Predictor variables included in the following models:

- Y1: K6 (Past Year)
- Y2: K6 (Past Year) + Age
- Y3: K6 (Past Year) + WHODAS
- Y4: K6 (Past Year) + WHODAS + Age
- Y5: K6 (Past Year) + WHODAS + Age + Past Year MDE
- Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts**
- M1: K6 (Past Month)
- M2: K6 (Past Month) + Age
- M3: K6 (Past Month) + WHODAS
- M4: K6 (Past Month) + WHODAS + Age
- M5: K6 (Past Month) + WHODAS + Age + Past Year MDE
- M6: K6 (Past Month) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Data and weights:

- Dataset = 2008A-2012 MHSS clinical data.
- Analysis weight = MHFAAWGT.

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

The following broad conclusions can be drawn from the ROC statistics.

SMI estimation:

- Models that do not include the WHODAS term tend to exhibit a higher bias and/or total error rate, suggesting that including this term is important to control bias and the total error rate.
- Excluding suicidal thoughts from the model appears to have a smaller effect on the bias and total error rate.
- Using past month K6 instead of past year K6 in the models does not seem to affect the bias or total error rate that much, suggesting that past month K6 could be used without substantially affecting the predictions if the past year version is unavailable.

AMI estimation:

- All of the models that include past month K6 show fairly large increases in bias, particularly those that do not include WHODAS, suggesting that the past month version of K6 in any model may lead to biased estimates of AMI.
- Models that include the past year K6 and WHODAS terms tend to exhibit reasonably low levels of bias and total error rate, and adding the age term (which is available in most surveys) improves those measures even further.

This page intentionally left blank

4. SMI and AMI National Estimates and Domain-Level Bias among NSDUH Models

4.1 Domain-Level Bias of Estimates among Parsimonious Models

As noted in Section 3.2 in Chapter 3, the cut point estimation method relies on being able to identify a cut point that at least approximately equalizes the weighted false-positive and false-negative rates to provide unbiased estimates of serious mental illness (SMI) (or any mental illness [AMI]), where bias is defined as the difference between the cut point estimate of SMI (or AMI) and the direct estimate of SMI (or AMI) as computed in the clinical sample. Even if this can be achieved, there is no guarantee that the same cut point will provide approximately unbiased estimates for various demographic and geographic domains. In fact, domain-level bias was a major consideration in the development of the 2012 prediction model, which led to the addition of the age variable in the model (Center for Behavioral Health Statistics and Quality [CBHSQ], 2014).

Domain-level bias was assessed for all 12 National Survey on Drug Use and Health (NSDUH) models described in [Table 3.1](#) (i.e., the 2012 model and 11 alternative models) for both SMI and AMI estimates across the following demographic and geographic domains:

- gender (male, female);
- age (18 to 25, 26 to 34, 35 to 49, 50 or older);
- race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic other, Hispanic);
- region (Northeast, Midwest, South, West);
- county type (large metro, small metro, nonmetro);
- received mental health services ("yes," "no");
- employment (full time, part time, unemployed, other);
- education (less than high school, high school graduate, some college, college graduate);
- poverty level (less than 100 percent, at least 100 percent but less than 200 percent, at least 200 percent); and
- health insurance ("yes," "no").

The domain-level bias results for the six models containing the past year version of the Kessler-6 (K6) scale for SMI and AMI are displayed in [Tables 4.1](#) and [4.2](#), respectively, and the results for the six models containing the past month version of the K6 scale for SMI and AMI are displayed in [Tables 4.3](#) and [4.4](#), respectively.

Table 4.1 Domain-Level Bias of SMI Cut Point Estimates Based on Models That Include K6 (Past Year) Term: 2008A-2012

Demographic Variable	Direct Estimate	Model = Y1 K6		Model = Y2 K6, Age		Model = Y3 K6, WHODAS		Model = Y4 K6, WHODAS, Age		Model = Y5 K6, WHODAS, Age, MDE		Model = Y6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Total	3.93	0.21	0.34	-0.21	0.33	0.00	0.30	0.05	0.29	0.00	0.29	-0.01	0.27
Gender													
Male	2.96	0.24	0.49	-0.01	0.49	-0.19	0.43	-0.07	0.43	0.02	0.41	0.08	0.42
Female	4.84	0.17	0.46	-0.41	0.45	0.18	0.41	0.17	0.40	-0.03	0.40	-0.10	0.34
Age													
18-25	3.77	2.91 ^b	0.75	0.20	0.64	1.81 ^b	0.69	0.26	0.61	0.28	0.64	-0.02	0.55
26-34	4.35	1.88 ^a	0.81	1.77 ^a	0.81	0.68	0.61	1.14	0.68	0.50	0.60	0.69	0.66
35-49	5.74	-1.59 ^a	0.71	-1.59 ^a	0.71	-1.67 ^b	0.61	-1.15 ^a	0.59	-1.05	0.58	-1.10	0.57
50+	2.74	-0.30	0.52	-0.30	0.52	0.12	0.48	0.28	0.48	0.33	0.47	0.37	0.42
Race/Ethnicity													
White, not Hispanic	4.43	-0.14	0.40	-0.56	0.39	-0.15	0.37	-0.02	0.37	-0.11	0.36	-0.17	0.33
Black, not Hispanic	3.28	0.25	0.77	-0.02	0.76	-0.15	0.60	-0.09	0.71	-0.21	0.58	-0.48	0.45
Other, not Hispanic	4.09	-0.09	1.51	-0.46	1.51	-0.71	1.14	-1.06	1.13	-1.33	1.13	-1.14	1.13
Hispanic	2.02	1.93	1.05	1.38	1.03	1.19	0.87	1.07	0.77	1.32	0.80	1.63	0.85
Region													
Northeast	2.80	-0.31	0.48	-0.63	0.47	-0.01	0.41	-0.24	0.39	-0.16	0.41	-0.04	0.39
Midwest	4.17	0.45	0.66	0.09	0.66	0.38	0.60	0.54	0.60	0.41	0.61	0.16	0.60
South	3.74	0.77	0.59	0.30	0.58	0.47	0.54	0.58	0.55	0.32	0.53	-0.13	0.52
West	5.04	-0.48	0.89	-0.98	0.88	-1.14	0.74	-1.03	0.70	-0.82	0.66	0.02	0.56
County Type													
Large Metro	3.78	-0.11	0.47	-0.45	0.47	-0.11	0.40	-0.04	0.40	-0.32	0.38	-0.34	0.38
Small Metro	4.15	0.90	0.64	0.45	0.63	0.38	0.60	0.48	0.59	0.45	0.57	0.73	0.49
Nonmetro	3.99	-0.10	0.68	-0.72	0.63	-0.37	0.59	-0.45	0.58	0.14	0.61	-0.38	0.59
Received Mental Health Services													
Yes	18.84	-4.70 ^b	1.67	-5.76 ^b	1.66	-2.45	1.51	-1.26	1.52	-0.99	1.46	-1.37	1.31
No	1.54	0.99 ^b	0.28	0.68 ^a	0.27	0.39	0.24	0.26	0.24	0.15	0.24	0.20	0.23
Employment													
Full Time	2.36	0.78 ^a	0.39	0.54	0.39	0.47	0.32	0.45	0.32	0.20	0.33	0.38	0.33
Part Time	4.34	1.39	0.76	0.35	0.73	0.46	0.63	0.03	0.58	0.91	0.60	0.39	0.59
Unemployed	5.64	1.15	1.46	0.71	1.45	-0.12	0.80	0.55	1.14	-0.53	0.79	0.03	0.75
Other	6.21	-1.60 ^a	0.81	-2.06 ^b	0.80	-1.04	0.79	-0.77	0.77	-0.68	0.74	-0.93	0.66
Education													
< High School	5.69	1.92	1.30	0.98	1.25	0.32	1.04	0.38	0.97	0.06	0.90	-0.11	0.90
High School Grad	4.05	0.40	0.73	0.02	0.73	-0.07	0.70	0.00	0.71	-0.04	0.69	0.01	0.59
Some College	4.14	0.33	0.49	-0.21	0.48	0.34	0.37	0.21	0.37	0.31	0.34	0.18	0.37
College Grad	2.88	-0.82	0.48	-0.96 ^a	0.48	-0.37	0.44	-0.18	0.44	-0.28	0.46	-0.16	0.46

See notes at end of table.

(continued)

Table 4.1 Domain-Level Bias of SMI Cut Point Estimates Based on Models That Include K6 (Past Year) Term: 2008A-2012 (continued)

Demographic Variable	Direct Estimate	Model = Y1 K6		Model = Y2 K6, Age		Model = Y3 K6, WHODAS		Model = Y4 K6, WHODAS, Age		Model = Y5 K6, WHODAS, Age, MDE		Model = Y6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Poverty Level¹													
< 100% Threshold	9.01	-0.63	1.42	-1.23	1.41	-1.20	1.37	-1.47	1.39	-1.73	1.30	-2.07	1.27
100% to 199% Threshold	5.61	1.00	0.98	0.19	0.95	-0.28	0.78	-0.34	0.79	-0.35	0.77	0.12	0.62
≥ 200% Threshold	2.59	0.11	0.33	-0.14	0.32	0.28	0.29	0.42	0.28	0.38	0.28	0.32	0.28
Health Insurance													
Yes	3.57	-0.23	0.34	-0.58	0.33	-0.04	0.31	-0.01	0.31	0.01	0.30	-0.16	0.27
No	5.71	2.39 ^a	1.12	1.62	1.08	0.21	0.84	0.34	0.87	-0.07	0.85	0.71	0.86

Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

^a Bias is statistically significant from zero at the 0.05 level.

^b Bias is statistically significant from zero at the 0.01 level.

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

NOTE: The predictor variables listed in the models are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicide (i.e., Suicidal Thoughts), and Past Year MDE. The model Y6 describes the 2012 model.

NOTE: Bias = Difference between cut point estimate from indicated model and direct estimate of SMI as computed in the clinical sample.

Dataset = 2008A-2012 MHSS clinical data.

Analysis weight = MHFAAWGT.

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

Table 4.2 Domain-Level Bias of AMI Cut Point Estimates Based on Models That Include K6 (Past Year) Term: 2008A-2012

Demographic Variable	Direct Estimate	Model = Y1 K6		Model = Y2 K6, Age		Model = Y3 K6, WHODAS		Model = Y4 K6, WHODAS, Age		Model = Y5 K6, WHODAS, Age, MDE		Model = Y6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Total	18.03	-0.56	1.17	0.26	1.17	0.13	1.11	0.00	1.11	0.01	1.11	-0.06	1.11
Gender													
Male	14.52	0.48	1.28	0.57	1.23	0.14	1.16	-0.26	1.14	-0.02	1.17	-0.02	1.18
Female	21.29	-1.54	1.92	-0.04	1.95	0.12	1.86	0.25	1.87	0.03	1.85	-0.11	1.84
Age													
18-25	21.20	7.35 ^b	2.39	0.22	2.37	6.17 ^b	2.36	0.49	2.11	0.45	2.15	-2.18	2.00
26-34	19.55	1.99	2.17	2.62	2.17	-0.19	1.88	0.43	1.94	0.39	1.88	0.35	1.90
35-49	20.53	-2.07	1.69	0.38	1.74	-0.74	1.58	0.38	1.58	-0.10	1.57	0.69	1.57
50+	14.84	-3.38	2.22	-0.72	2.23	-1.31	2.14	-0.55	2.17	-0.24	2.18	0.05	2.20
Race/Ethnicity													
White, not Hispanic	18.66	-1.15	1.05	0.19	1.02	0.30	0.94	0.18	0.94	0.29	0.95	0.18	0.96
Black, not Hispanic	15.01	1.10	2.67	1.13	2.69	-0.32	2.48	0.42	2.56	-0.67	2.46	-0.57	2.47
Other, not Hispanic	15.10	5.62	3.48	6.04	3.80	1.96	3.27	2.23	3.35	1.91	3.27	0.96	3.25
Hispanic	18.93	-2.15	5.84	-2.94	5.84	-1.20	5.77	-2.23	5.69	-1.72	5.72	-1.31	5.69
Region													
Northeast	19.14	-4.91	4.28	-4.64	4.30	-6.17	4.07	-6.01	4.07	-5.93	4.09	-5.63	4.10
Midwest	17.26	-0.10	1.42	0.88	1.45	0.84	1.32	0.33	1.30	0.18	1.30	-0.17	1.31
South	16.84	1.36	1.39	1.75	1.48	3.32 ^b	1.25	3.09 ^a	1.26	3.43 ^b	1.28	3.55 ^b	1.27
West	19.69	-0.05	2.07	1.83	1.84	0.15	1.92	0.32	1.93	-0.14	1.88	-0.57	1.87
County Type													
Large Metro	19.54	-1.36	2.04	-1.23	2.03	-1.57	1.94	-1.74	1.93	-2.03	1.92	-1.97	1.91
Small Metro	16.50	-0.54	1.19	0.49	1.18	1.98	1.12	1.76	1.12	2.40 ^a	1.17	2.33	1.20
Nonmetro	16.19	1.87	1.89	4.48 ^a	1.91	1.92	1.62	2.15	1.67	1.86	1.67	1.38	1.70
Received Mental Health Services													
Yes	52.96	-7.33 ^a	3.46	-2.86	3.18	0.25	3.02	0.62	3.03	0.87	2.99	1.75	2.90
No	12.41	0.52	1.25	0.75	1.27	0.10	1.20	-0.10	1.20	-0.14	1.20	-0.36	1.20
Employment													
Full Time	14.65	0.56	1.18	1.72	1.21	-0.53	1.03	-0.42	1.04	-0.56	1.02	-0.22	1.02
Part Time	20.28	1.80	2.47	0.91	2.41	2.92	2.23	3.13	2.23	2.51	2.23	2.12	2.30
Unemployed	21.21	3.13	3.81	3.31	3.44	0.21	3.68	-1.00	3.67	-1.38	3.62	-2.29	3.65
Other	22.42	-4.52	3.11	-3.40	3.12	0.03	3.15	-0.43	3.14	0.21	3.18	-0.29	3.15
Education													
< High School	26.05	-2.14	6.47	-2.86	6.50	-0.34	6.54	-1.61	6.45	-1.45	6.46	-2.17	6.43
High School Grad	17.88	-0.06	1.84	0.40	1.82	-0.70	1.55	-0.32	1.58	-0.65	1.61	-0.40	1.57
Some College	15.58	2.32	1.37	2.99 ^a	1.36	2.79 ^a	1.30	2.50	1.28	2.46	1.26	2.32	1.24
College Grad	16.93	-2.96	1.52	-1.00	1.49	-1.29	1.36	-1.24	1.37	-0.96	1.36	-0.99	1.39

See notes at end of table.

(continued)

Table 4.2 Domain-Level Bias of AMI Cut Point Estimates Based on Models That Include K6 (Past Year) Term: 2008A-2012 (continued)

Demographic Variable	Direct Estimate	Model = Y1 K6		Model = Y2 K6, Age		Model = Y3 K6, WHODAS		Model = Y4 K6, WHODAS, Age		Model = Y5 K6, WHODAS, Age, MDE		Model = Y6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Poverty Level¹													
< 100% Threshold	25.34	3.34	2.68	2.80	2.67	4.34	2.58	3.77	2.58	3.22	2.52	2.20	2.54
100% to 199% Threshold	24.48	-2.13	4.79	-1.84	4.77	-0.92	4.69	-1.02	4.68	-1.36	4.67	-1.22	4.70
≥ 200% Threshold	15.05	-0.97	0.98	0.30	0.99	-0.53	0.88	-0.50	0.88	-0.30	0.89	-0.18	0.88
Health Insurance													
Yes	17.11	-1.00	1.31	0.03	1.29	0.10	1.24	0.06	1.24	0.08	1.24	-0.12	1.24
No	22.57	1.60	2.60	1.40	2.80	0.25	2.54	-0.26	2.52	-0.38	2.52	0.19	2.45

AMI = any mental illness; Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error; WHODAS = World Health Organization Disability Assessment Schedule.

^a Bias is statistically significant from zero at the 0.05 level.

^b Bias is statistically significant from zero at the 0.01 level.

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

NOTE: The predictor variables listed in the models are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicide (i.e., Suicidal Thoughts), and Past Year MDE. The model Y6 describes the 2012 model.

NOTE: Bias = Difference between cut point estimate from indicated model and direct estimate of AMI as computed in the clinical sample.

Dataset = 2008A-2012 MHSS clinical data.

Analysis weight = MHFAAWGT.

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

Table 4.3 Domain-Level Bias of SMI Cut Point Estimates Based on Models That Include K6 (Past Month) Term: 2008A-2012

Demographic Variable	Direct Estimate	Model = M1 K6		Model = M2 K6, Age		Model = M3 K6, WHODAS		Model = M4 K6, WHODAS, Age		Model = M5 K6, WHODAS, Age, MDE		Model = M6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Total	3.93	-0.47	0.33	0.03	0.33	-0.01	0.31	0.02	0.31	0.00	0.28	0.01	0.27
Gender													
Male	2.96	0.10	0.49	0.55	0.49	0.16	0.45	0.20	0.45	0.06	0.41	0.07	0.42
Female	4.84	-0.99 ^a	0.45	-0.45	0.45	-0.18	0.43	-0.15	0.42	-0.06	0.38	-0.05	0.34
Age													
18-25	3.77	2.40 ^b	0.83	1.37	0.80	0.63	0.66	-0.69	0.56	0.17	0.63	0.14	0.55
26-34	4.35	-1.01	0.71	-0.25	0.74	0.60	0.69	0.85	0.69	0.00	0.56	0.49	0.63
35-49	5.74	-2.16 ^b	0.68	-1.12	0.68	-1.35 ^a	0.61	-1.00	0.62	-1.05	0.58	-1.06	0.58
50+	2.74	-0.23	0.53	0.37	0.53	0.33	0.52	0.54	0.52	0.57	0.46	0.41	0.42
Race/Ethnicity													
White, not Hispanic	4.43	-0.75	0.41	-0.15	0.41	-0.32	0.39	-0.24	0.38	0.01	0.36	-0.13	0.33
Black, not Hispanic	3.28	0.41	0.82	0.73	0.86	0.13	0.77	0.19	0.73	-0.23	0.56	-0.08	0.49
Other, not Hispanic	4.09	-1.67	1.41	-1.28	1.45	-0.97	1.19	-0.94	1.20	-1.54	1.15	-1.62	1.08
Hispanic	2.02	0.76	0.83	0.96	0.86	1.75 ^a	0.86	1.52	0.84	0.89	0.75	1.50	0.83
Region													
Northeast	2.80	-0.07	0.50	0.13	0.50	0.29	0.48	0.45	0.49	-0.04	0.40	0.28	0.42
Midwest	4.17	-0.10	0.68	0.62	0.70	0.26	0.64	0.34	0.64	0.35	0.60	0.12	0.59
South	3.74	-0.20	0.61	0.38	0.62	0.39	0.58	0.38	0.56	0.14	0.52	-0.10	0.52
West	5.04	-1.64 ^a	0.78	-1.24	0.77	-1.23	0.73	-1.32	0.72	-0.54	0.65	-0.18	0.55
County Type													
Large Metro	3.78	-0.37	0.47	0.13	0.48	-0.14	0.43	-0.19	0.42	-0.06	0.38	-0.26	0.38
Small Metro	4.15	-0.51	0.62	-0.21	0.62	0.36	0.61	0.45	0.61	-0.05	0.56	0.52	0.48
Nonmetro	3.99	-0.69	0.66	0.16	0.64	-0.31	0.63	-0.16	0.65	0.28	0.57	-0.12	0.60
Received Mental Health Services													
Yes	18.84	-7.61 ^b	1.61	-5.27 ^b	1.59	-2.49	1.60	-2.28	1.60	-1.53	1.41	-0.81	1.32
No	1.54	0.68 ^a	0.28	0.88 ^b	0.29	0.35	0.25	0.35	0.24	0.24	0.23	0.14	0.23
Employment													
Full Time	2.36	-0.27	0.36	0.18	0.37	0.20	0.34	0.25	0.34	0.04	0.32	0.26	0.33
Part Time	4.34	-0.15	0.77	0.34	0.81	-0.09	0.64	-0.33	0.61	0.71	0.61	0.61	0.61
Unemployed	5.64	0.37	1.51	0.48	1.47	0.40	1.21	0.33	1.22	-0.10	0.81	-0.32	0.67
Other	6.21	-1.16	0.82	-0.49	0.81	-0.45	0.80	-0.32	0.80	-0.37	0.72	-0.65	0.66
Education													
< High School	5.69	1.89	1.18	2.45 ^a	1.18	1.26	1.07	1.15	1.06	-0.15	0.81	-0.12	0.88
High School Grad	4.05	0.03	0.80	0.71	0.80	0.17	0.72	0.12	0.71	0.08	0.70	0.16	0.59
Some College	4.14	-1.28 ^b	0.42	-0.86 ^a	0.42	-0.25	0.39	-0.09	0.40	0.10	0.33	0.18	0.35
College Grad	2.88	-1.21 ^b	0.46	-0.85	0.48	-0.52	0.48	-0.47	0.48	-0.11	0.45	-0.25	0.47

See notes at end of table.

(continued)

Table 4.3 Domain-Level Bias of SMI Cut Point Estimates Based on Models That Include K6 (Past Month) Term: 2008A-2012 (continued)

Demographic Variable	Direct Estimate	Model = M1 K6		Model = M2 K6, Age		Model = M3 K6, WHODAS		Model = M4 K6, WHODAS, Age		Model = M5 K6, WHODAS, Age, MDE		Model = M6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Poverty Level¹													
< 100% Threshold	9.01	-1.46	1.47	-0.76	1.47	-0.03	1.47	-0.28	1.46	-2.06	1.26	-1.78	1.24
100% to 199% Threshold	5.61	-0.49	0.84	0.34	0.84	-0.59	0.79	-0.42	0.80	-0.14	0.69	-0.06	0.59
≥ 200% Threshold	2.59	-0.28	0.34	0.09	0.34	0.13	0.30	0.19	0.30	0.39	0.28	0.35	0.28
Health Insurance													
Yes	3.57	-0.65	0.34	-0.16	0.34	-0.15	0.32	-0.13	0.32	0.05	0.30	-0.13	0.28
No	5.71	0.44	1.07	0.95	1.07	0.65	0.94	0.72	0.94	-0.27	0.81	0.70	0.84

Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

^a Bias is statistically significant from zero at the 0.05 level.

^b Bias is statistically significant from zero at the 0.01 level.

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

NOTE: The predictor variables listed in the models are defined in exactly the same way as in the 2012 model (see Chapter 2): WHODAS, Age, Suicide (i.e., Suicidal Thoughts), and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead.

NOTE: Bias = Difference between cut point estimate from indicated model and direct estimate of SMI as computed in the clinical sample.

Dataset = 2008A-2012 MHSS clinical data.

Analysis weight = MHFAAWGT.

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

Table 4.4 Domain-Level Bias of AMI Cut Point Estimates Based on Models That Include K6 (Past Month) Term: 2008A-2012

Demographic Variable	Direct Estimate	Model = M1 K6		Model = M2 K6, Age		Model = M3 K6, WHODAS		Model = M4 K6, WHODAS, Age		Model = M5 K6, WHODAS, Age, MDE		Model = M6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Total	18.03	-2.69 ^a	1.19	-3.51 ^b	1.18	-0.98	1.10	0.96	1.13	-0.89	1.10	-1.08	1.10
Gender													
Male	14.52	-1.27	1.32	-1.96	1.31	-1.02	1.13	0.30	1.17	-0.77	1.15	-0.80	1.17
Female	21.29	-4.01 ^a	1.93	-4.95 ^b	1.90	-0.94	1.84	1.57	1.90	-1.00	1.84	-1.34	1.82
Age													
18-25	21.20	3.57	2.41	-1.97	2.19	4.21	2.27	1.23	2.18	0.59	2.17	-2.20	2.01
26-34	19.55	-1.64	2.43	-1.64	2.43	-1.05	1.86	-0.39	1.88	-0.64	1.87	-0.77	1.88
35-49	20.53	-4.24 ^a	1.98	-4.24 ^a	1.98	-2.41	1.53	2.55	1.59	-1.82	1.53	-1.19	1.53
50+	14.84	-4.32 ^a	2.15	-4.32 ^a	2.15	-1.87	2.14	0.43	2.23	-0.94	2.17	-0.75	2.18
Race/Ethnicity													
White, not Hispanic	18.66	-4.00 ^b	1.00	-4.54 ^b	1.00	-1.06	0.93	1.60	0.96	-0.63	0.94	-0.98	0.94
Black, not Hispanic	15.01	0.79	2.73	-1.15	2.60	-0.40	2.48	0.44	2.53	-0.48	2.46	-0.91	2.47
Other, not Hispanic	15.10	-0.35	3.49	-1.22	3.43	0.73	3.17	2.89	3.87	0.41	3.21	-0.05	3.14
Hispanic	18.93	-0.51	6.31	-1.70	6.24	-1.87	5.73	-2.56	5.69	-3.05	5.66	-2.17	5.66
Region													
Northeast	19.14	-7.04	4.13	-7.47	4.11	-6.73	4.05	-5.08	4.09	-6.07	4.08	-6.00	4.09
Midwest	17.26	-0.82	1.48	-1.72	1.48	0.06	1.30	1.39	1.32	-0.45	1.28	-0.87	1.29
South	16.84	-0.25	1.50	-1.00	1.49	2.21	1.23	4.04 ^b	1.27	2.39	1.25	2.40	1.23
West	19.69	-4.47	2.29	-5.67 ^a	2.26	-1.79	1.88	1.20	2.12	-1.77	1.87	-2.30	1.85
County Type													
Large Metro	19.54	-4.86 ^a	2.00	-5.67 ^b	1.98	-2.83	1.91	-1.06	1.94	-2.95	1.90	-3.06	1.89
Small Metro	16.50	-1.55	1.48	-2.35	1.47	0.81	1.08	3.02 ^a	1.25	1.39	1.14	1.29	1.17
Nonmetro	16.19	1.93	1.77	1.09	1.76	1.44	1.69	3.38	1.79	1.25	1.69	0.65	1.70
Received Mental Health Services													
Yes	52.96	-13.92 ^b	2.94	-15.16 ^b	2.90	-1.17	3.00	3.03	3.04	-0.45	2.98	0.15	2.88
No	12.41	-0.90	1.30	-1.65	1.29	-0.95	1.18	0.62	1.22	-0.97	1.19	-1.29	1.19
Employment													
Full Time	14.65	-2.90 ^a	1.21	-3.37 ^b	1.21	-1.84	1.01	0.12	1.05	-1.57	1.00	-1.57	1.00
Part Time	20.28	-3.87	2.32	-4.89 ^a	2.32	1.50	2.20	3.08	2.30	0.88	2.18	0.64	2.26
Unemployed	21.21	3.60	4.16	-0.54	3.72	-0.84	3.56	1.35	3.89	-1.40	3.60	-2.07	3.67
Other	22.42	-3.20	3.23	-3.80	3.21	-0.56	3.14	1.41	3.22	-0.34	3.17	-0.75	3.13
Education													
< High School	26.05	0.09	6.91	-0.85	6.85	-1.00	6.49	0.23	6.67	-2.13	6.41	-2.99	6.37
High School Grad	17.88	-1.21	1.84	-1.79	1.85	-1.12	1.54	0.02	1.57	-1.18	1.59	-1.06	1.54
Some College	15.58	-1.02	1.35	-2.65 ^a	1.27	1.23	1.25	3.30 ^a	1.35	1.27	1.24	0.81	1.21
College Grad	16.93	-6.77 ^b	1.49	-7.04 ^b	1.49	-2.81 ^a	1.36	0.05	1.37	-2.03	1.36	-1.99	1.39

See notes at end of table.

(continued)

Table 4.4 Domain-Level Bias of AMI Cut Point Estimates Based on Models That Include K6 (Past Month) Term: 2008A-2012 (continued)

Demographic Variable	Direct Estimate	Model = M1 K6		Model = M2 K6, Age		Model = M3 K6, WHODAS		Model = M4 K6, WHODAS, Age		Model = M5 K6, WHODAS, Age, MDE		Model = M6 K6, WHODAS, Age, MDE, Suicide	
		Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias	Bias	SE Bias
Poverty Level¹													
< 100% Threshold	25.34	1.83	2.65	-0.37	2.57	3.41	2.44	6.21 ^a	2.79	1.90	2.40	0.59	2.38
100% to 199% Threshold	24.48	-0.93	5.15	-2.29	5.10	-1.67	4.67	-1.24	4.68	-2.02	4.65	-2.41	4.66
≥ 200% Threshold	15.05	-4.04 ^b	0.93	-4.45 ^b	0.93	-1.75 ^a	0.87	0.48	0.89	-1.23	0.88	-1.03	0.87
Health Insurance													
Yes	17.11	-3.56 ^b	1.27	-4.31 ^b	1.25	-1.01	1.22	1.15	1.26	-0.67	1.23	-0.87	1.23
No	22.57	1.62	3.23	0.46	3.23	-0.81	2.49	-0.03	2.58	-2.00	2.44	-2.13	2.37

AMI = any mental illness; Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error; WHODAS = World Health Organization Disability Assessment Schedule.

^a Bias is statistically significant from zero at the 0.05 level.

^b Bias is statistically significant from zero at the 0.01 level.

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

NOTE: The predictor variables listed in the models are defined in exactly the same way as in the 2012 model (see Chapter 2): WHODAS, Age, Suicide (i.e., Suicidal Thoughts), and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead.

NOTE: Bias = Difference between cut point estimate from indicated model and direct estimate of AMI as computed in the clinical sample.

Dataset = 2008A-2012 MHSS clinical data.

Analysis weight = MHFAAWGT.

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

Table 4.1 indicates eight cases of significant (at the 0.05 or 0.01 level) domain-level bias with respect to SMI estimates for the model containing only past year K6; the number of significantly biased cases reduces to six if an age variable is added to the model. The model containing the past year K6 and WHODAS scales indicates bias in only two age groups; when age is added to this model, bias persists for one of these domains. The 2012 model or the model that includes all of the terms in the 2012 model excluding suicidal thoughts shows no significant bias for any domain.

Table 4.2 indicates far fewer cases (one or two) of significant domain-level bias with respect to AMI estimates for any of the models containing the past year K6 scale. The 2012 model and the model that contains the past year K6, WHODAS, and age terms indicate only a single significantly biased case (i.e., Region = South has significant bias in both models).

Table 4.3 displays a similar pattern to Table 4.1 in terms of domain-level bias with respect to SMI estimates for models containing the past month K6 scale. In fact, fewer cases of domain-level bias are in Table 4.3 than in Table 4.1. This finding appears to be similar to the finding in Section 3.2 that, in terms of SMI estimation, there appears to be little loss in accuracy if the past month version of the K6 scale is used instead of the past year version. However, with respect to the AMI estimation models that contain the past month K6 scale but not the WHODAS scale, Table 4.4 shows many cases with significant domain-level bias and significant bias overall. Adding the WHODAS scale to the model substantially improves the domain-level bias profile, and the 2012 model or the model that includes all of the terms in the 2012 model excluding suicidal thoughts shows no significant bias for any domain.

4.2 Comparisons of SMI and AMI National Estimates among NSDUH Models

Although the previous section examined the differences in domain-level bias across the different models, this section discusses the actual estimates of SMI and AMI produced by these different models based on the combined 2008A-2012 WHODAS data. National estimates of SMI and AMI from all 12 models are displayed in Tables 4.5 and 4.6, respectively (shown after this section's text ends for reader convenience). Even though for each of the 11 alternative models estimates are shown for the entire sample, any discussion of differences is actually between the clinical sample and the model-based estimates based on the clinical sample. This is because appropriate comparisons between national estimates obtained from model-based predictions and direct estimates derived from clinical determinations need to be conducted using data in which each respondent possesses both a model-based prediction of SMI and AMI and a direct estimate of those measures; in other words, appropriate comparisons can only be made using the clinical sample data, which satisfies this requirement. A consequence of this is that these comparisons are subject to larger design-based sampling errors because of the substantially smaller sample size of the clinical sample. The differences between estimates from the various models and direct estimates were tested using the clinical sample data, and the results of the tests are indicated in Tables 4.5 and 4.6. The purpose of the tests was to assess the extent to which estimates based on the models compare with the direct estimates from the clinical sample.

Table 4.5 indicates that with respect to SMI estimates, models that do not include the WHODAS scale showed many significant differences at the domain level and in particular

among domains related to age and received mental health services. Adding the WHODAS scale to the model removed significant differences across all domains except for some related to age. The further addition of age and past year MDE to the model removed all significant differences at the domain level.

Table 4.6 indicates that with respect to AMI estimates, some of the model-based estimates appeared to differ fairly substantially from the direct estimates without being significantly different. This may be because the model-based estimates are derived from the adult NSDUH data (i.e., where design-based sampling error is smaller), but the tests between the model-based estimates and the direct estimates were based on the clinical sample (i.e., where design-based sampling error is larger). Even so, models that include only the past month K6 scale with or without age showed significant differences even at the overall level. All other models show significant differences for at most three domains in a scattered fashion, and models that include the past month K6, WHODAS, age, and past year MDE terms showed no significant differences at the domain level.

Table 4.5 NSDUH Estimates of SMI Based on Models of Varying Parsimony, by Demographic Domain: 2008A-2012

Demographic Variable	Direct Estimate	Model											
		Y1	Y2	Y3	Y4	Y5	Y6	M1	M2	M3	M4	M5	M6
Total	3.93	4.28	3.73	3.89	3.80	4.02	3.88	3.73	4.16	3.86	3.89	3.84	3.77
Gender													
Male	2.96	3.30	2.88	2.69	2.62	2.74	2.88	3.20	3.52	2.78	2.78	2.69	2.85
Female	4.84	5.19	4.52	5.01	4.89	5.21	4.81	4.23 ^a	4.75	4.86	4.92	4.92	4.62
Age													
18-25	3.77	7.76 ^b	4.23	5.77 ^b	3.98	4.19	3.86	6.13 ^b	4.66	5.10	3.69	3.97	3.79
26-34	4.35	5.87 ^a	5.68 ^a	5.04	5.14	4.91	4.98	4.36	4.98	4.61	4.84	4.54	4.72
35-49	5.74	4.51 ^a	4.51 ^a	4.49 ^b	4.79 ^a	5.14	4.99	3.94 ^b	4.90	4.49 ^a	4.83	4.78	4.82
50+	2.74	2.32	2.32	2.43	2.59	2.90	2.75	2.52	3.21	2.74	2.99	2.93	2.72
Race/Ethnicity													
White, not Hispanic	4.43	4.21	3.69	4.20	4.13	4.40	4.23	3.58	4.07	4.10	4.15	4.19	4.10
Black, not Hispanic	3.28	4.33	3.71	3.00	2.92	3.16	3.09	4.57	4.82	3.15	3.25	3.12	3.04
Other, not Hispanic	4.09	3.74	3.11	3.46	3.29	3.20	3.19	2.99	3.31	3.80	3.66	2.98	2.95
Hispanic	2.02	4.83	4.21	3.35	3.16	3.28	3.14	4.11	4.46	3.32 ^a	3.27	3.19	3.16
Region													
Northeast	2.80	3.96	3.41	3.50	3.33	3.62	3.55	3.39	3.79	3.47	3.50	3.47	3.50
Midwest	4.17	4.44	3.84	4.05	3.96	4.33	4.23	3.80	4.24	4.00	4.06	4.05	4.10
South	3.74	4.29	3.79	3.88	3.83	4.01	3.70	3.92	4.36	3.92	3.94	3.80	3.53
West	5.04	4.37	3.78	4.08	3.96	4.05	4.09	3.62 ^a	4.08	3.93	3.95	4.00	4.03
County Type													
Large Metro	3.78	4.08	3.53	3.58	3.49	3.66	3.62	3.53	3.89	3.57	3.61	3.52	3.51
Small Metro	4.15	4.51	3.94	4.14	4.00	4.36	4.13	3.92	4.38	4.05	4.01	4.15	3.97
Nonmetro	3.99	4.50	3.99	4.44	4.42	4.54	4.25	4.04	4.64	4.44	4.57	4.32	4.20
Received Mental Health Services													
Yes	18.84	16.13 ^b	14.69 ^b	17.71	17.76	19.17	18.52	13.31 ^b	15.08 ^b	17.29	17.85	18.46	17.94
No	1.54	2.39 ^b	1.98 ^a	1.69	1.57	1.60	1.54	2.19 ^a	2.41 ^b	1.71	1.66	1.51	1.50
Employment													
Full Time	2.36	3.22 ^a	2.85	2.74	2.73	2.82	2.78	2.41	2.77	2.50	2.56	2.56	2.60
Part Time	4.34	4.88	3.76	4.40	3.95	4.27	4.05	3.72	3.86	4.06	3.85	4.01	3.92
Unemployed	5.64	7.98	6.70	6.34	5.93	6.34	6.41	7.45	7.78	6.53	6.19	6.20	6.15
Other	6.21	5.11 ^a	4.67 ^b	5.17	5.15	5.52	5.20	5.30	6.01	5.59	5.76	5.53	5.25
Education													
< High School	5.69	6.38	5.62	4.27	4.12	4.18	3.98	6.79	7.35 ^a	4.58	4.54	4.30	4.14
High School Grad	4.05	4.73	4.09	4.26	4.09	4.22	4.13	4.36	4.83	4.27	4.27	4.08	4.03
Some College	4.14	4.53	3.82	4.63	4.41	4.75	4.43	3.50 ^b	3.83 ^a	4.43	4.44	4.42	4.36
College Grad	2.88	2.49	2.28 ^a	2.64	2.77	3.06	3.05	1.68 ^b	2.09	2.53	2.65	2.83	2.76

See notes at end of table.

(continued)

Table 4.5 NSDUH Estimates of SMI Based on Models of Varying Parsimony, by Demographic Domain: 2008A-2012 (continued)

Demographic Variable	Direct Estimate	Model											
		Y1	Y2	Y3	Y4	Y5	Y6	M1	M2	M3	M4	M5	M6
Poverty Level¹													
< 100% Threshold	9.01	8.93	7.84	7.49	7.17	7.35	6.96	8.71	9.30	7.74	7.72	7.34	7.06
100% to 199% Threshold	5.61	5.68	5.01	4.86	4.80	4.85	4.76	5.14	5.79	4.91	4.90	4.74	4.67
≥ 200% Threshold	2.59	2.92	2.54	2.88	2.83	3.11	3.00	2.32	2.67	2.77	2.83	2.88	2.85
Health Insurance													
Yes	3.57	3.84	3.36	3.63	3.57	3.79	3.64	3.31	3.73	3.59	3.64	3.62	3.50
No	5.71	6.55 ^a	5.64	5.18	4.93	5.18	5.13	5.87	6.38	5.17	5.11	4.95	5.16

Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the direct estimate as computed in the clinical sample is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the direct estimate as computed in the clinical sample is statistically significant from zero at the 0.01 level.

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

Predictor variables included in the following models:

Y1: K6 (Past Year)

Y2: K6 (Past Year) + Age

Y3: K6 (Past Year) + WHODAS

Y4: K6 (Past Year) + WHODAS + Age

Y5: K6 (Past Year) + WHODAS + Age + Past Year MDE

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

M3: K6 (Past Month) + WHODAS

M4: K6 (Past Month) + WHODAS + Age

M5: K6 (Past Month) + WHODAS + Age + Past Year MDE

M6: K6 (Past Month) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

The predictor variables above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Datasets = 2008A-2012 adult NSDUH data for model-based estimates, and 2008A-2012 MHSS clinical data for direct estimates and difference tests.

Analysis weight = ANALWT_A/5 for model-based estimates, and MHFAAWGT for direct estimates and difference tests.

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

Table 4.6 NSDUH Estimates of AMI Based on Models of Varying Parsimony, by Demographic Domain: 2008A-2012

Demographic Variable	Direct Estimate	Model											
		Y1	Y2	Y3	Y4	Y5	Y6	M1	M2	M3	M4	M5	M6
Total	18.03	17.51	16.44	17.79	17.83	18.94	18.05	15.77 ^a	15.00 ^b	17.12	19.02	20.16	17.47
Gender													
Male	14.52	14.38	13.35	14.06	14.01	14.96	14.33	13.98	13.23	13.80	15.24	16.16	14.02
Female	21.29	20.42	19.32	21.26	21.38	22.65	21.52	17.44 ^a	16.65 ^b	20.22	22.53	23.90	20.68
Age													
18-25	21.20	28.19 ^b	21.38	26.92 ^b	21.84	22.72	18.57	24.94	19.72	25.47	21.92	23.29	18.70
26-34	19.55	22.15	21.77	21.47	21.92	23.37	21.79	18.68	18.68	20.11	21.85	23.48	20.88
35-49	20.53	18.32	18.32	18.22	19.59	20.61	20.36	16.08 ^a	16.08 ^a	17.46	20.64	21.80	19.33
50+	14.84	11.52	11.52	12.95	13.77	14.89	15.00	11.28 ^a	11.28 ^a	12.87	15.90	16.79	14.57
Race/Ethnicity													
White, not Hispanic	18.66	17.72	16.78	18.47	18.60	19.77	18.99	15.38 ^b	14.72 ^b	17.68	19.93	21.11	18.27
Black, not Hispanic	15.01	17.59	16.24	16.78	16.67	17.75	16.65	17.53	16.56	16.55	17.81	19.04	16.66
Other, not Hispanic	15.10	17.48	16.17	16.81	16.94	17.99	16.77	16.90	15.88	16.45	17.66	18.70	16.21
Hispanic	18.93	16.43	15.11	15.79	15.53	16.38	15.29	15.67	14.66	15.22	16.25	17.24	14.86
Region													
Northeast	19.14	17.31	16.27	17.25	17.37	18.40	17.54	15.33	14.56	16.61	18.45	19.47	16.93
Midwest	17.26	17.94	16.89	18.02	18.10	19.14	18.38	16.02	15.20	17.35	19.15	20.23	17.71
South	16.84	17.05	16.00	17.62 ^b	17.64 ^a	18.70 ^b	17.83 ^b	15.73	15.02	17.01	18.78 ^b	19.94	17.35
West	19.69	17.97	16.87	18.27	18.25	19.56	18.52	15.94	15.13 ^a	17.50	19.72	21.01	17.84
County Type													
Large Metro	19.54	16.98	15.88	17.07	17.12	18.34	17.38	15.25 ^a	14.46 ^b	16.43	18.27	19.49	16.81
Small Metro	16.50	18.04	16.96	18.64	18.60	19.65	18.81	16.11	15.33	17.86	19.88 ^a	21.02	18.17
Nonmetro	16.19	18.25	17.34 ^a	18.53	18.71	19.58	18.81	16.83	16.14	18.01	19.82	20.79	18.31
Received Mental Health Services													
Yes	52.96	49.05 ^a	47.86	51.93	52.67	55.25	53.87	41.56 ^b	40.66 ^b	50.22	53.95	56.73	52.39
No	12.41	12.47	11.42	12.34	12.27	13.15	12.33	11.64	10.90	11.84	13.44	14.33	11.89
Employment													
Full Time	14.65	14.91	14.23	14.51	14.97	15.94	15.23	12.53 ^a	12.07 ^b	13.61	15.71	16.86	14.43
Part Time	20.28	21.08	18.76	21.04	20.12	21.47	19.67	17.89	16.17 ^a	20.19	20.90	22.27	19.22
Unemployed	21.21	27.78	25.11	25.28	24.18	25.45	23.66	27.46	25.49	24.67	25.01	26.58	23.32
Other	22.42	18.32	17.51	20.47	20.46	21.66	21.09	18.12	17.49	20.28	22.68	23.64	20.75
Education													
< High School	26.05	20.59	19.29	19.82	19.90	20.72	19.62	21.86	20.78	20.28	21.70	22.60	19.62
High School Grad	17.88	17.74	16.51	17.64	17.44	18.56	17.81	16.90	15.93	17.18	18.67	19.78	17.41
Some College	15.58	18.96	17.58 ^a	19.71 ^a	19.45	20.61	19.45	16.17	15.24 ^a	18.65	20.41 ^a	21.65	18.68
College Grad	16.93	14.34	13.86	15.15	15.70	16.91	16.22	11.05 ^b	10.80 ^b	14.04 ^a	16.73	17.95	15.30

See notes at end of table.

(continued)

Table 4.6 NSDUH Estimates of AMI Based on Models of Varying Parsimony, by Demographic Domain: 2008A-2012 (continued)

Demographic Variable	Direct Estimate	Model											
		Y1	Y2	Y3	Y4	Y5	Y6	M1	M2	M3	M4	M5	M6
Poverty Level¹													
< 100% Threshold	25.34	27.76	25.62	26.79	26.06	27.14	25.56	27.36	25.87	26.41	26.94 ^a	28.26	25.15
100% to 199% Threshold	24.48	21.21	19.91	21.09	21.08	22.18	21.06	20.20	19.27	20.51	22.11	23.33	20.59
≥ 200% Threshold	15.05	14.31	13.58	14.93	15.20	16.32	15.68	12.09 ^b	11.56 ^b	14.19 ^a	16.50	17.59	15.02
Health Insurance													
Yes	17.11	16.39	15.43	16.99	17.09	18.21	17.44	14.54 ^b	13.85 ^b	16.34	18.46	19.57	16.85
No	22.57	23.23	21.64	21.88	21.66	22.73	21.22	22.01	20.86	21.11	21.88	23.24	20.68

AMI = any mental illness; Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error, WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the direct estimate as computed in the clinical sample is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the direct estimate as computed in the clinical sample is statistically significant from zero at the 0.01 level.

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

Predictor variables included in the following models:

Y1: K6 (Past Year)

Y2: K6 (Past Year) + Age

Y3: K6 (Past Year) + WHODAS

Y4: K6 (Past Year) + WHODAS + Age

Y5: K6 (Past Year) + WHODAS + Age + Past Year MDE

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

M3: K6 (Past Month) + WHODAS

M4: K6 (Past Month) + WHODAS + Age

M5: K6 (Past Month) + WHODAS + Age + Past Year MDE

M6: K6 (Past Month) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

The predictor variables above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Datasets = 2008A-2012 adult NSDUH data for model-based estimates, and 2008A-2012 MHSS clinical data for direct estimates and difference tests.

Analysis weight = ANALWT_A/5 for model-based estimates, and MHFAAWGT for direct estimates and difference tests.

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

This page intentionally left blank

5. Analyzing SMI or AMI in Conjunction with Model Predictor Variables

The goal of the minimum-biased cut point methodology described in Chapter 2 is to predict the status ("yes" or "no") of serious mental illness (SMI) and any mental illness (AMI) for every adult responding to the main survey interview of the National Survey on Drug Use and Health (NSDUH) to provide (as close as possible) unbiased prevalence estimates of SMI and AMI at the national level and within various demographic and geographic domains.

However, as discussed in Chapter 1, analyses of mental illness variables in conjunction with model predictor variables, or variables that are strongly associated with these variables, present particular problems because of the structural relationship between SMI or AMI and the model predictor variables induced by the model itself. Examples of such analyses include domain analyses in which the domains are defined by one or more model predictor variables (e.g., the percentage of adults with SMI by whether or not they had suicidal thoughts or past year major depressive episode [MDE]) or regression analyses involving a model predictor variable (e.g., past year MDE regressed on SMI). The analyses that follow provide more information about the problems associated with the analysis of mental illness variables in conjunction with the model predictor and related variables.

Of the five alternative models that include past year Kessler-6 (K6) examined in this report (i.e., models Y1 to Y5), all exclude suicidal thoughts as a predictor variable, and four of them exclude past year MDE as a predictor variable (i.e., models Y1 to Y4). The variables associated with suicidal thoughts or past year MDE are often used in conjunction with SMI or AMI in analyses. Therefore, for comparison purposes, this chapter investigates the relationship between SMI or AMI and past year MDE and suicidal thoughts using three models, all of which include past year K6, the World Health Organization Disability Assessment Schedule (WHODAS), and age as "base" predictor variables. The three models include the following: (1) Y4 (excludes past year MDE and suicidal thoughts), (2) Y5 (excludes suicidal thoughts), and (3) Y6 (the 2012 model that includes both past year MDE and suicidal thoughts). Note that similar analyses using models that include past month K6 instead of past year K6 (i.e., models M4 to M6) were not conducted because the differences between the two sets of models involve only the K6 variables; therefore, the models including past month K6 are unlikely to yield any further relevant information related to these investigative analyses.

The results in [Table 5.1](#) demonstrate the association between SMI and AMI and related variables associated with the MDE and suicide⁴ predictive variables in the three models (i.e., Y4

⁴ In creating the SMI and AMI predicted values under the models that included suicidal thoughts as a predictor variable, missing values for past year suicidal thoughts were treated as zeroes. For creating the domains in this chapter's tables, they are treated as missing (i.e., not allocated to a domain).

Table 5.1 Comparing SMI and AMI Prevalence Estimates for MDE and Suicidal Thought Domains Using Models Excluding or Including Past Year MDE or Suicidal Thoughts as Predictor Variables

Domain	Direct Estimate: MHSS Clinical Sample	NSDUH Estimate: Model = Y4 K6, WHODAS, Age	NSDUH Estimate: Model = Y5 K6, WHODAS, Age, MDE	NSDUH Estimate: Model = Y6 K6, WHODAS, Age, MDE, Suicide
SMI				
Past Year MDE				
Yes	34.1	37.2	54.1 ^a	49.1 ^a
No	1.7	1.3	0.5 ^a	0.7 ^a
Lifetime MDE				
Yes	21.1	22.6	29.2 ^a	27.0 ^a
No	1.3	0.9	0.3 ^a	0.5 ^a
Had Suicidal Thoughts in the Past Year				
Yes	41.8	32.6 ^a	36.5	56.7 ^a
No	2.5	2.6	2.7	1.8 ^a
Made Suicidal Plans in the Past Year				
Yes	57.9	44.2 ^a	49.0	66.2 ^a
No	3.3	3.3	3.5	3.2
Had Suicidal Attempt in the Past Year				
Yes	47.2	48.6	41.8	59.3
No	3.8	3.5	3.8	3.6
AMI				
Past Year MDE				
Yes	72.4	84.2 ^a	99.9 ^a	98.7 ^a
No	14.1	13.0	13.1	12.3
Lifetime MDE				
Yes	52.4	62.9 ^a	71.1 ^a	69.4 ^a
No	12.9	11.1	11.2	10.4 ^a
Had Suicidal Thoughts in the Past Year				
Yes	68.3	74.5 ^a	76.4 ^a	99.6 ^a
No	16.1	15.6	16.7	14.9
Made Suicidal Plans in the Past Year				
Yes	79.6	81.3	82.6	99.5 ^a
No	17.4	17.1	18.2	17.2
Had Suicidal Attempt in the Past Year				
Yes	92.4	77.1	78.1	99.4
No	17.8	17.5	18.6	17.6

AMI = any mental illness; CBHSQ = Center for Behavioral Health Statistics and Quality; K6 = Kessler-6; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; 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 sample using the bias measure described in Section 7.3.3 of CBHSQ (2014) is significantly different from zero at the .05 level.

NOTE: The predictor variables listed in the models are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicide (i.e., Suicidal Thoughts), and Past Year MDE.

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

to Y6). Of the three models, two did not include suicidal thoughts as a predictor variable (i.e., models Y4 and Y5) and one did not include past year MDE as a predictor variable (i.e., model Y4). The table includes as domain variables not only the variables used in the predictive model, but also associated variables such as lifetime MDE and whether the adult had suicidal plans or attempts in the past year.

The results in [Table 5.1](#) show that SMI and AMI cut point estimates for these domains were often not close to the direct estimates from the Mental Health Surveillance Study (MHSS) clinical sample. Estimates of SMI and AMI were *overestimated* among respondents *with* past year MDE or lifetime MDE for models that included past year MDE as a predictor variable (i.e., models Y5 and Y6). Additionally, estimates of SMI were *underestimated* among respondents *without* past year MDE or lifetime MDE for the same models. Similarly, estimates of SMI and AMI were *overestimated* among respondents *with* thoughts or plans of suicide for the model that included suicidal thoughts as a predictor variable (i.e., model Y6).

The observed overestimation (and underestimation in the converse situation in some cases) 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 (or AMI) to exhibit a larger correlation with that variable than exists between the variable and the clinical diagnosis of SMI (or AMI).

To summarize, the tendency for estimates based on the 2012 model to be biased upward also exists in domains having definitions overlapping past year MDE or suicidal thoughts, such as lifetime MDE or suicidal *plans* (e.g., an adult with a suicidal plan also had a suicidal thought). 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 MDE or suicidal thoughts. Clearly, *the mental illness prevalence estimates derived from the 2012 model should not be used for these domains or for domains with overlapping definitions.*

Analysts are sometimes interested not only in SMI prevalence within different domains, but also in the attributes of adults within the domain defined by having SMI. [Table 5.2](#) looks at the prevalence of suicidal thoughts, plans, and attempts in such a domain.

[Table 5.2](#) displays the prevalence estimates of suicidal thoughts, plans, and attempts among adults with SMI as estimated by the clinical sample using the clinical determination of SMI and as estimated in the NSDUH main survey sample using the three models described above (i.e., models Y4, Y5, and Y6) to predict SMI. Estimates from all three models were biased for suicidal thoughts (i.e., the direct estimate was statistically significantly different from the three modeled estimates). However, note that the estimate for suicidal thoughts using model Y6 (54.8 percent) was substantially *larger* than the direct estimate (39.9 percent); on the other hand, the estimates from the other two models that excluded suicidal thoughts as a predictor variable were both smaller than the direct estimate.

Table 5.2 Comparing Prevalence Estimates of Suicidal Thoughts, Plans, and Attempts in the Past Year among Adults with SMI Using Models Excluding or Including Suicidal Thoughts as a Predictor Variable

Domain	Direct Estimate: MHSS Clinical Sample	NSDUH Estimate: Model = Y4 K6, WHODAS, Age	NSDUH Estimate: Model = Y5 K6, WHODAS, Age, MDE	NSDUH Estimate: Model = Y6 K6, WHODAS, Age, MDE, Suicide
Proportion with SMI Who Had Suicidal Thoughts	39.9	33.0 ^a	34.1 ^a	54.8 ^a
Proportion with SMI Who Planned Suicide	15.7	12.6	12.9	18.0
Proportion with SMI Who Attempted Suicide	4.1	5.1	5.1	7.5

CBHSQ = Center for Behavioral Health Statistics and Quality; K6 = Kessler-6; MDE = major depressive episode; MHSS = Mental Health Surveillance Study; 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 sample using the bias measure described in Section 7.3.3 of CBHSQ (2014) is significantly different from zero at the .05 level.

NOTE: The predictor variables listed in the models are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicide (i.e., Suicidal Thoughts), and Past Year MDE.

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

6. Comparisons of SMI and AMI National Estimates between NSDUH and Other Data Sources

One of the goals of this study was to see whether any suitable alternative models could be identified and applied to other data sources (where some, but not all, of the predictor variables of the 2012 model would be available) for the purpose of generating reasonably comparable mental illness estimates. This goal was evaluated following the identification of a selected group of recent health studies that included at least some of the predictor variables of the 2012 model.

The other health studies examined in this chapter include the Behavioral Risk Factor Surveillance System (BRFSS), the National Health Interview Survey (NHIS), and the Medical Expenditure Panel Survey (MEPS). Table 6.1 illustrates which serious mental illness (SMI) predictor variables are included in each of these other health studies; note that only the past month Kessler-6 (K6) and age predictor variables are included in each of the three other studies.

Table 6.1 Predictor Variables Included in NSDUH and Other Studies

Predictor Variable	NSDUH	BRFSS	NHIS	MEPS
K6 (Past Month)	X	X	X	X
Age	X	X	X	X
K6 (Past Year)	X			
WHODAS	X			
Past Year MDE	X			
Suicidal Thoughts	X			

BRFSS = Behavioral Risk Factor Surveillance System; K6 = Kessler-6; MDE = major depressive episode; MEPS = Medical Expenditure Panel Survey; NHIS = National Health Interview Survey; NSDUH = National Survey on Drug Use and Health; WHODAS = World Health Organization Disability Assessment Schedule.

NOTE: An "X" indicates which predictor variable is included in each other study and NSDUH.

Even if estimates generated by different sources are comparable for the same model, the relationship between the survey design and the estimates cannot be discounted. Therefore, some of the methodological characteristics of NSDUH, BRFSS, NHIS, and MEPS data sources are first briefly described to illustrate differences among them. This description includes a list of the information relevant to the various SMI prediction models discussed in Chapter 3 that is available and collected from these data sources. Then SMI and any mental illness (AMI) estimates are compared between those derived from NSDUH and those derived from the three other data sources.

6.1 Short Descriptions of NSDUH and Other Data Sources

6.1.1 National Survey on Drug Use and Health (NSDUH)

The National Survey on Drug Use and Health (NSDUH) is conducted annually and is sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA). It is the primary source of statistical information on the use of illegal drugs by the civilian,

noninstitutionalized population of the United States aged 12 or older and also includes assessments of mental problems, mental health service use, and other health-related behaviors. NSDUH is a nationally representative survey that uses a state-based design with an independent, multistage area probability sample within each state and the District of Columbia to produce national, state, and substate estimates. The annual national target sample size for all respondents aged 12 or older is 67,500, and this includes a target sample size for adult respondents aged 18 or older of 45,000. Because it asks potentially sensitive questions, NSDUH uses an audio computer-assisted self-interviewing (ACASI) administration mode. For further details about NSDUH, see the *2014 National Survey on Drug Use and Health: Methodological summary and definitions* (Center for Behavioral Health Statistics and Quality, 2015).

6.1.2 Behavioral Risk Factor Surveillance System (BRFSS)

BRFSS is a state-based system of health surveys that collects information on health risk behaviors, clinical preventive practices, and health care access and use primarily related to chronic diseases and injury. BRFSS has technical and methodological assistance from the Centers for Disease Control and Prevention (CDC). States conduct monthly telephone surveys of noninstitutionalized adults aged 18 or older using random-digit dialing (RDD) methods. BRFSS started in 1984 and has, since 1994, collected data from all 50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands, and Guam using a computer-assisted telephone interviewing (CATI) design. The BRFSS design allows states to add optional modules; one such module, incorporated in the 2007 BRFSS, focuses on mental illness and "stigma" (i.e., perceived prejudice and discrimination). In 2007, this optional module was administered by 35 states, the District of Columbia, and Puerto Rico, in which data from 202,114 respondents were obtained. This optional module contains the past month K6 scale and questions on attitudes toward persons with mental illness. For further details about BRFSS, see the BRFSS website (CDC, 2014).

6.1.3 National Health Interview Survey (NHIS)

NHIS is sponsored by the National Center for Health Statistics (NCHS) and is a continuous nationally representative sample survey that collects data using personal household interviews through an interviewer-administered computer-assisted personal interviewing (CAPI) system. NHIS data have been collected since 1957, and in 2012, there were 21,781 respondents aged 18 or older (NCHS, Division of Health Interview Statistics, 2013). The survey provides national estimates of a broad range of health measures, including health status and health care access. Since 1997, NHIS has included the past month K6 scale in the questionnaire. For further details about NHIS, see the NHIS website (CDC, 2015).

6.1.4 Medical Expenditure Panel Survey (MEPS)

MEPS began in 1996 and is sponsored by the Agency for Healthcare Research and Quality (AHRQ). It is a set of large-scale surveys of families and individuals, their medical providers (e.g., doctors, hospitals, pharmacies), and employers across the United States. MEPS collects data on the specific health services that Americans use, how frequently the services are used, the cost of these services, and how they are paid for, as well as data on the cost, scope, and breadth of health insurance held by and available to U.S. workers. MEPS currently contains a Household Component, which uses a CAPI system to survey individuals in households; this

information is supplemented by data from their medical providers. The MEPS sample is drawn from NHIS respondents; thus, NHIS and MEPS panel data can be linked for analysis. MEPS uses mail-back self-administered paper-and-pencil questionnaires (the self-administered questionnaire [SAQ]) for questions that may be unreliable if answered by a proxy during the MEPS core household interview. The SAQ includes the past month K6 scale and is administered to all household respondents aged 18 or older. For further details about MEPS, see the MEPS website (AHRQ, n.d.).

6.2 Comparisons of SMI and AMI Estimates between NSDUH and Other Data Sources

Each of the three non-NSDUH data sources collects information that corresponds to only two of the models listed in [Table 3.1](#), namely, M1 (i.e., with one covariate: past month K6) and M2 (i.e., with two covariates: past month K6 and age). Therefore, for each other data source, the following four comparisons with NSDUH were conducted for both SMI and AMI estimates:

- M1 (NSDUH) versus M1 (other data source),
- M2 (NSDUH) versus M2 (other data source),
- Y6⁵ (NSDUH) versus M1 (other data source), and
- Y6 (NSDUH) versus M2 (other data source).

The first two compare estimates between NSDUH and the other data source using the same alternative model in each case, and the next two compare estimates between the current NSDUH model against each of the two models available for the other data source.

[Tables 6.2 to 6.4](#) display the results of the four comparisons with respect to SMI for each of the three other data sources, and [Tables 6.5 to 6.7](#) display similar results with respect to AMI.⁶ The tables include comparisons among demographic and geographic domains that are as close as possible to those listed in [Section 4.1](#), but an exact match was not always possible. For example, BRFSS had no equivalent variable to describe poverty level, and there was ambiguity in how full-time employment and part-time employment were defined (so these categories were collapsed). Also, NHIS had no equivalent variable to describe county type. Moreover, MEPS had no equivalent variables to describe county type or the receipt of mental health services, and there was ambiguity in how full-time employment and part-time employment were defined (so these categories were collapsed).

[Tables 6.2 and 6.3](#) indicate that, with respect to SMI, most of the differences among the four comparisons with regard to BRFSS were significantly different, and almost all of the same differences with regard to NHIS were significantly different. For example, the overall SMI estimate based on the 2012 NSDUH model M1 was 4.12 percent, but corresponding SMI estimates from the 2012 BRFSS and 2012 NHIS were 3.18 and 2.43 percent, respectively. In other words, overall SMI estimates from BRFSS and particularly NHIS based on model M1

⁵ Model Y6 is the 2012 model used to provide national estimates of SMI and AMI based on adult NSDUH data.

⁶ For the reader's convenience, [Tables 6.2 to 6.7](#) are presented together after this text discussion regarding them ends on the next page.

were much lower than the corresponding NSDUH estimate. This pattern was similar for the other comparisons and for the domains within all comparisons.

On the other hand, [Table 6.4](#) indicates that, with respect to SMI, some of the differences among the four comparisons with regard to MEPS were not significantly different. For example, the overall SMI estimate based on the 2010 NSDUH model Y6 was 4.07 percent, and the corresponding SMI estimate from the 2010 MEPS was 4.10 percent (this difference was not significant). However, there were still many differences across the domains and across the different comparisons that were significantly different.

[Tables 6.5](#) to [6.7](#) indicate that, with respect to AMI, the differences overall and among almost every domain for all of the comparisons with regard to all three other data sources were significant. For example, the overall AMI estimate based on the 2012 NSDUH model Y6 was 18.59 percent, but corresponding AMI estimates from the 2012 BRFSS model M1 and the 2012 NHIS model M1 were much lower at 11.85 and 8.94 percent, respectively. The overall AMI estimate based on the 2010 NSDUH model Y6 was 18.06 percent, but the corresponding AMI estimate from the 2010 MEPS was still substantially lower at 13.79 percent.

In summary, estimates of SMI and AMI derived from the three other data sources appeared to be substantially lower than those based on NSDUH, particularly those derived from NHIS, but less so for those derived from MEPS. The different methodological characteristics of all of these data sources may account for a large part of the observed differences in the estimates.

Table 6.2 Comparison of Estimates of SMI between Selected NSDUH and BRFSS Models: 2012

Demographic Variable	BRFSS Model = M1	NSDUH Model = M1	NSDUH Model = Y6	BRFSS Model = M2	NSDUH Model = M2	NSDUH Model = Y6
Total	3.18	4.12 ^b	4.05 ^b	3.64	4.59 ^b	4.05 ^b
Gender						
Male	2.82	3.52 ^b	3.16	3.29	3.94 ^b	3.16
Female	3.51	4.68 ^b	4.88 ^b	3.97	5.19 ^b	4.88 ^b
Age						
18-25	2.62	6.31 ^b	4.11 ^b	1.79	4.77 ^b	4.11 ^b
26-34	2.84	4.81 ^b	5.17 ^b	3.26	5.47 ^b	5.17 ^b
35-49	3.96	4.15	5.17 ^b	4.81	5.23	5.17
50+	3.04	3.12	2.96	3.72	3.83	2.96 ^b
Race/Ethnicity						
White, not Hispanic	2.68	3.77 ^b	4.21 ^b	3.10	4.27 ^b	4.21 ^b
Black, not Hispanic	3.32	5.84 ^b	3.41	3.97	6.24 ^b	3.41
Other, not Hispanic	3.62	3.14	2.92	4.24	3.49	2.92 ^a
Hispanic	5.06	4.85	4.40	5.29	5.26	4.40
Region						
Northeast	2.72	3.71 ^b	3.65 ^b	3.30	4.09 ^b	3.65
Midwest	2.55	3.85 ^b	3.93 ^b	2.99	4.44 ^b	3.93 ^b
South	4.58	4.35	4.05	4.87	4.89	4.05
West	3.10	4.33 ^b	4.48 ^b	3.67	4.64 ^b	4.48 ^b
County Type						
Large Metro	3.13	3.99 ^b	3.84 ^b	3.55	4.43 ^b	3.84
Small Metro	3.64	4.28	4.43 ^a	4.11	4.76	4.43
Nonmetro	3.65	4.24 ^a	4.05	4.37	4.80	4.05
Received Mental Health Services						
Yes	14.79	13.21 ^a	17.59 ^b	17.02	15.07 ^a	17.59
No	1.93	2.58 ^b	1.75	2.21	2.80 ^b	1.75 ^b
Employment						
Full or Part Time	1.43	2.82 ^b	2.99 ^b	1.61	3.18 ^b	2.99 ^b
Unemployed	7.07	9.12 ^b	7.82	7.89	9.26	7.82
Other	4.96	5.92 ^b	5.58 ^a	5.76	6.67 ^b	5.58
Education						
< High School	7.76	8.05	4.67 ^b	8.62	8.76	4.67 ^b
High School Grad	3.26	4.94 ^b	4.43 ^b	3.83	5.43 ^b	4.43 ^a
Some College	2.81	3.64 ^b	4.38 ^b	3.21	4.05 ^b	4.38 ^b
College Grad	1.14	1.77 ^b	3.06 ^b	1.35	2.15 ^b	3.06 ^b
Health Insurance						
Yes	2.69	3.61 ^b	3.66 ^b	3.11	4.06 ^b	3.66 ^b
No	5.59	6.67 ^a	6.03	6.27	7.28	6.03

BRFSS = Behavioral Risk Factor Surveillance System; Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the estimate from the BRFSS model in the column to the left is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the estimate from the BRFSS model in the column to the left is statistically significant from zero at the 0.01 level.

Predictor variables included in the following models:

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Domains: BRFSS had no equivalent variable to describe poverty level, and there was ambiguity in how full-time employment and part-time employment were defined (so these categories were collapsed).

Datasets = 2012 adult NSDUH data, and 2012 adult BRFSS public use data.

Analysis weight = ANALWT for NSDUH, and BRFSS analysis weights.

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

Table 6.3 Comparison of Estimates of SMI between Selected NSDUH and NHIS Models: 2012

Demographic Variable	NHIS Model = M1	NSDUH Model = M1	NSDUH Model = Y6	NHIS Model = M2	NSDUH Model = M2	NSDUH Model = Y6
Total	2.43	4.12 ^b	4.05 ^b	2.81	4.59 ^b	4.05 ^b
Gender						
Male	1.82	3.52 ^b	3.16 ^b	2.10	3.94 ^b	3.16 ^b
Female	3.00	4.68 ^b	4.88 ^b	3.46	5.19 ^b	4.88 ^b
Age						
18-25	1.59	6.31 ^b	4.11 ^b	1.12	4.77 ^b	4.11 ^b
26-34	2.39	4.81 ^b	5.17 ^b	2.61	5.47 ^b	5.17 ^b
35-49	2.77	4.15 ^b	5.17 ^b	3.39	5.23 ^b	5.17 ^b
50+	2.52	3.12 ^a	2.96 ^a	3.09	3.83 ^b	2.96
Race/Ethnicity						
White, not Hispanic	2.38	3.77 ^b	4.21 ^b	2.79	4.27 ^b	4.21 ^b
Black, not Hispanic	2.62	5.84 ^b	3.41 ^a	3.03	6.24 ^b	3.41
Other, not Hispanic	1.88	3.14 ^b	2.92 ^a	2.11	3.49 ^b	2.92
Hispanic	2.69	4.85 ^b	4.40 ^b	2.99	5.26 ^b	4.40 ^b
Region						
Northeast	2.27	3.71 ^b	3.65 ^b	2.52	4.09 ^b	3.65 ^b
Midwest	2.26	3.85 ^b	3.93 ^b	2.78	4.44 ^b	3.93 ^b
South	2.68	4.35 ^b	4.05 ^b	3.09	4.89 ^b	4.05 ^b
West	2.31	4.33 ^b	4.48 ^b	2.61	4.64 ^b	4.48 ^b
Received Mental Health Services						
Yes	12.76	13.21	17.59 ^b	14.34	15.07	17.59 ^b
No	1.54	2.58 ^b	1.75 ^a	1.82	2.80 ^b	1.75
Employment						
Full Time	0.89	2.52 ^b	2.73 ^b	1.05	2.97 ^b	2.73 ^b
Part Time	2.00	3.88 ^b	3.91 ^b	2.04	3.95 ^b	3.91 ^b
Unemployed	4.93	9.12 ^b	7.82 ^b	5.10	9.26 ^b	7.82 ^b
Other	4.41	5.92 ^b	5.58 ^b	5.32	6.67 ^b	5.58
Education						
< High School	5.10	8.05 ^b	4.67	5.50	8.76 ^b	4.67 ^a
High School Grad	2.83	4.94 ^b	4.43 ^b	3.34	5.43 ^b	4.43 ^b
Some College	2.20	3.64 ^b	4.38 ^b	2.68	4.05 ^b	4.38 ^b
College Grad	0.81	1.77 ^b	3.06 ^b	0.95	2.15 ^b	3.06 ^b
Poverty Level¹						
< 100% Threshold	6.54	9.31 ^b	7.23	7.44	10.04 ^b	7.23
100-199% Threshold	3.65	5.37 ^b	5.16 ^b	4.45	5.99 ^b	5.16 ^a
≥ 200% Threshold	1.24	2.47 ^b	2.94 ^b	1.43	2.85 ^b	2.94 ^b
Health Insurance						
Yes	2.16	3.61 ^b	3.66 ^b	2.56	4.06 ^b	3.66 ^b
No	3.76	6.67 ^b	6.03 ^b	4.05	7.28 ^b	6.03 ^b

Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; NHIS = National Health Interview Survey; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the estimate from the NHIS model in the column to the left is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the estimate from the NHIS model in the column to the left is statistically significant from zero at the 0.01 level.

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

Predictor variables included in the following models:

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Domains: NHIS had no equivalent variable to describe county type.

Datasets = 2012 adult NSDUH data, and 2012 adult NHIS public use data.

Analysis weight = ANALWT for NSDUH, and NHIS analysis weights.

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

Table 6.4 Comparison of Estimates of SMI between Selected NSDUH and MEPS Models: 2010

Demographic Variable	MEPS Model = M1	NSDUH Model = M1	NSDUH Model = Y6	MEPS Model = M2	NSDUH Model = M2	NSDUH Model = Y6
Total	4.10	3.72 ^a	4.07	4.72	4.21 ^b	4.07 ^b
Gender						
Male	3.52	3.18	3.01 ^a	4.15	3.47 ^b	3.01 ^b
Female	4.65	4.24 ^a	5.05	5.24	4.91	5.05
Age						
18-25	2.96	6.19 ^b	3.92 ^b	2.29	4.59 ^b	3.92 ^b
26-34	3.06	4.20 ^b	5.09 ^b	3.32	4.74 ^b	5.09 ^b
35-49	4.82	3.82 ^b	5.25	5.85	4.86 ^b	5.25
50+	4.44	2.61 ^b	2.97 ^b	5.36	3.46 ^b	2.97 ^b
Race/Ethnicity						
White, not Hispanic	4.01	3.53 ^b	4.32	4.61	4.06 ^b	4.32
Black, not Hispanic	4.43	4.61	3.83	5.19	5.04	3.83 ^b
Other, not Hispanic	3.68	2.75	3.60	4.34	2.88 ^a	3.60
Hispanic	4.48	4.37	3.23 ^b	5.02	4.89	3.23 ^b
Region						
Northeast	3.48	3.95	3.90	4.12	4.65	3.90
Midwest	4.23	4.08	4.54	4.91	4.34	4.54
South	4.55	3.77 ^b	3.70 ^b	5.19	4.18 ^b	3.70 ^b
West	3.77	3.14 ^a	4.33	4.25	3.78	4.33
Employment						
Full or Part Time	1.77	2.41 ^b	3.06 ^b	2.05	2.72 ^b	3.06 ^b
Unemployed	2.80	3.61 ^b	4.10 ^b	3.44	3.67	4.10 ^a
Other	7.91	5.61 ^b	5.45 ^b	9.03	6.51 ^b	5.45 ^b
Education						
< High School	8.24	5.92 ^b	3.40 ^b	8.83	6.65 ^b	3.40 ^b
High School Grad	4.60	4.68	4.68	5.32	5.25	4.68
Some College	3.36	3.36	4.56 ^b	4.09	3.56 ^a	4.56
College Grad	1.87	1.88	3.32 ^b	2.31	2.40	3.32 ^b
Poverty Level¹						
< 100% Threshold	10.58	8.75 ^b	6.97 ^b	11.92	9.25 ^b	6.97 ^b
100-199% Threshold	5.70	4.84 ^a	4.69 ^b	6.45	5.55 ^a	4.69 ^b
≥ 200% Threshold	2.53	2.39	3.30 ^b	2.97	2.81	3.30
Health Insurance						
Yes	3.81	3.41 ^b	3.89	4.39	3.92 ^b	3.89 ^b
No	5.24	5.32	4.96	5.96	5.70	4.96 ^a

Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; MEPS = Medical Expenditure Panel Survey; SE = standard error; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the estimate from the MEPS model in the column to the left is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the estimate from the MEPS model in the column to the left is statistically significant from zero at the 0.01 level.

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

Predictor variables included in the following models:

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Domains: MEPS had no equivalent variables to describe county type or receipt of mental health services, and there was ambiguity in how full-time employment and part-time employment was defined (so these categories were collapsed).

Datasets = 2010 adult NSDUH data, and 2010 adult MEPS public use data.

Analysis weight = ANALWT for NSDUH, and MEPS analysis weights.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2010.

Table 6.5 Comparison of Estimates of AMI between Selected NSDUH and BRFSS Models: 2012

Demographic Variable	BRFSS Model = M1	NSDUH Model = M1	NSDUH Model = Y6	BRFSS Model = M2	NSDUH Model = M2	NSDUH Model = Y6
Total	11.85	16.13 ^b	18.59 ^b	11.30	15.37 ^b	18.59 ^b
Gender						
Male	10.52	14.47 ^b	14.89 ^b	9.94	13.74 ^b	14.89 ^b
Female	13.10	17.66 ^b	22.02 ^b	12.58	16.89 ^b	22.02 ^b
Age						
18-25	14.36	25.16 ^b	19.56 ^b	10.52	20.05 ^b	19.56 ^b
26-34	10.46	19.32 ^b	22.87 ^b	10.46	19.32 ^b	22.87 ^b
35-49	12.97	15.96 ^b	20.21 ^b	12.97	15.96 ^b	20.21 ^b
50+	10.92	12.02 ^b	15.76 ^b	10.92	12.02 ^b	15.76 ^b
Race/Ethnicity						
White, not Hispanic	10.15	15.61 ^b	19.29 ^b	9.66	14.97 ^b	19.29 ^b
Black, not Hispanic	14.04	19.10 ^b	18.55 ^b	13.19	18.27 ^b	18.55 ^b
Other, not Hispanic	14.46	15.90	16.99 ^a	13.79	14.84	16.99 ^b
Hispanic	16.24	16.25	16.27	15.68	15.17	16.27
Region						
Northeast	11.72	15.27 ^b	17.99 ^b	11.21	14.60 ^b	17.99 ^b
Midwest	9.90	16.05 ^b	18.16 ^b	9.34	15.24 ^b	18.16 ^b
South	14.56	16.66 ^b	18.69 ^b	13.95	15.97 ^b	18.69 ^b
West	11.80	16.01 ^b	19.29 ^b	11.30	15.15 ^b	19.29 ^b
County Type						
Large Metro	12.03	15.66 ^b	17.99 ^b	11.34	14.86 ^b	17.99 ^b
Small Metro	12.83	16.90 ^b	19.61 ^b	12.28	16.20 ^b	19.61 ^b
Nonmetro	13.24	16.22 ^b	18.65 ^b	12.72	15.54 ^b	18.65 ^b
Received Mental Health Services						
Yes	39.98	40.53	52.37 ^b	39.13	39.66	52.37 ^b
No	9.30	11.95 ^b	12.83 ^b	8.73	11.22 ^b	12.83 ^b
Employment						
Full or Part Time	7.86	13.68 ^b	16.23 ^b	7.46	12.98 ^b	16.23 ^b
Unemployed	23.62	29.23 ^b	25.50	22.37	27.12 ^b	25.50 ^a
Other	15.28	18.80 ^b	22.25 ^b	14.65	18.19 ^b	22.25 ^b
Education						
< High School	23.32	23.21	21.86	22.34	22.18	21.86
High School Grad	12.97	17.78 ^b	18.69 ^b	12.28	16.73 ^b	18.69 ^b
Some College	10.75	15.63 ^b	19.67 ^b	10.11	14.85 ^b	19.67 ^b
College Grad	6.00	11.36 ^b	15.86 ^b	5.93	11.08 ^b	15.86 ^b
Health Insurance						
Yes	10.25	14.94 ^b	17.85 ^b	9.84	14.25 ^b	17.85 ^b
No	19.30	22.10 ^b	22.31 ^b	18.39	21.10 ^b	22.31 ^b

AMI = any mental illness; BRFSS = Behavioral Risk Factor Surveillance System; Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; SE = standard error; WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the estimate from the BRFSS model in the column to the left is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the estimate from the BRFSS model in the column to the left is statistically significant from zero at the 0.01 level.

Predictor variables included in the following models:

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Domains: BRFSS had no equivalent variable to describe poverty level, and there was ambiguity in how full-time and part-time employment were defined (so these categories were collapsed).

Datasets = 2012 adult NSDUH data, and 2012 adult BRFSS public use data.

Analysis weight = ANALWT for NSDUH, and BRFSS analysis weights.

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

Table 6.6 Comparison of Estimates of AMI between Selected NSDUH and NHIS Models: 2012

Demographic Variable	NHIS Model = M1	NSDUH Model = M1	NSDUH Model = Y6	NHIS Model = M2	NSDUH Model = M2	NSDUH Model = Y6
Total	8.94	16.13 ^b	18.59 ^b	8.61	15.37 ^b	18.59 ^b
Gender						
Male	7.41	14.47 ^b	14.89 ^b	6.97	13.74 ^b	14.89 ^b
Female	10.37	17.66 ^b	22.02 ^b	10.14	16.89 ^b	22.02 ^b
Age						
18-25	9.10	25.16 ^b	19.56 ^b	6.82	20.05 ^b	19.56 ^b
26-34	8.45	19.32 ^b	22.87 ^b	8.45	19.32 ^b	22.87 ^b
35-49	9.50	15.96 ^b	20.21 ^b	9.50	15.96 ^b	20.21 ^b
50+	8.74	12.02 ^b	15.76 ^b	8.74	12.02 ^b	15.76 ^b
Race/Ethnicity						
White, not Hispanic	8.83	15.61 ^b	19.29 ^b	8.54	14.97 ^b	19.29 ^b
Black, not Hispanic	8.80	19.10 ^b	18.55 ^b	8.44	18.27 ^b	18.55 ^b
Other, not Hispanic	7.15	15.90 ^b	16.99 ^b	6.76	14.84 ^b	16.99 ^b
Hispanic	10.30	16.25 ^b	16.27 ^b	9.85	15.17 ^b	16.27 ^b
Region						
Northeast	7.94	15.27 ^b	17.99 ^b	7.66	14.60 ^b	17.99 ^b
Midwest	9.00	16.05 ^b	18.16 ^b	8.57	15.24 ^b	18.16 ^b
South	9.41	16.66 ^b	18.69 ^b	9.19	15.97 ^b	18.69 ^b
West	8.95	16.01 ^b	19.29 ^b	8.50	15.15 ^b	19.29 ^b
Received Mental Health Services						
Yes	35.15	40.53 ^b	52.37 ^b	33.75	39.66 ^b	52.37 ^b
No	6.70	11.95 ^b	12.83 ^b	6.45	11.22 ^b	12.83 ^b
Employment						
Full Time	4.90	12.58 ^b	15.24 ^b	4.76	12.12 ^b	15.24 ^b
Part Time	8.11	17.64 ^b	19.79 ^b	7.45	16.07 ^b	19.79 ^b
Unemployed	16.43	29.23 ^b	25.50 ^b	15.29	27.12 ^b	25.50 ^b
Other	13.84	18.80 ^b	22.25 ^b	13.53	18.19 ^b	22.25 ^b
Education						
< High School	15.92	23.21 ^b	21.86 ^b	15.31	22.18 ^b	21.86 ^b
High School Grad	9.52	17.78 ^b	18.69 ^b	9.11	16.73 ^b	18.69 ^b
Some College	9.39	15.63 ^b	19.67 ^b	9.05	14.85 ^b	19.67 ^b
College Grad	4.46	11.36 ^b	15.86 ^b	4.36	11.08 ^b	15.86 ^b
Poverty Level¹						
< 100% Threshold	18.67	27.53 ^b	26.80 ^b	18.07	26.24 ^b	26.80 ^b
100-199% Threshold	13.23	19.87 ^b	21.80 ^b	12.37	19.00 ^b	21.80 ^b
≥ 200% Threshold	5.85	12.13 ^b	15.61 ^b	5.69	11.61 ^b	15.61 ^b
Health Insurance						
Yes	8.20	14.94 ^b	17.85 ^b	7.96	14.25 ^b	17.85 ^b
No	12.59	22.10 ^b	22.31 ^b	11.81	21.10 ^b	22.31 ^b

AMI = any mental illness; Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; NHIS = National Health Interview Survey; SE = standard error; WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the estimate from the NHIS model in the column to the left is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the estimate from the NHIS model in the column to the left is statistically significant from zero at the 0.01 level.

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

Predictor variables included in the following models:

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Domains: NHIS had no equivalent variable to describe county type.

Datasets = 2012 adult NSDUH data, and 2012 adult NHIS public use data.

Analysis weight = ANALWT for NSDUH, and NHIS analysis weights.

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

Table 6.7 Comparison of Estimates of AMI between Selected NSDUH and MEPS Models: 2010

Demographic Variable	MEPS Model = M1	NSDUH Model = M1	NSDUH Model = Y6	MEPS Model = M2	NSDUH Model = M2	NSDUH Model = Y6
Total	13.79	16.15 ^b	18.06 ^b	13.50	15.40 ^b	18.06 ^b
Gender						
Male	12.25	14.48 ^b	14.83 ^b	11.97	13.74 ^b	14.83 ^b
Female	15.24	17.72 ^b	21.09 ^b	14.94	16.95 ^b	21.09 ^b
Age						
18-25	11.16	24.64 ^b	18.01 ^b	9.15	19.57 ^b	18.01 ^b
26-34	13.22	18.38 ^b	21.57 ^b	13.22	18.38 ^b	21.57 ^b
35-49	14.29	16.16 ^b	20.58 ^b	14.29	16.16 ^b	20.58 ^b
50+	14.58	12.31 ^b	15.13	14.58	12.31 ^b	15.13
Race/Ethnicity						
White, not Hispanic	13.56	15.35 ^b	19.04 ^b	13.28	14.69 ^b	19.04 ^b
Black, not Hispanic	14.29	18.61 ^b	16.95 ^b	14.04	17.58 ^b	16.95 ^b
Other, not Hispanic	13.91	18.32 ^b	15.94	13.43	17.43 ^b	15.94 ^a
Hispanic	14.44	17.00 ^b	15.18	14.14	16.09 ^a	15.18
Region						
Northeast	13.36	16.73 ^b	18.64 ^b	13.14	15.89 ^b	18.64 ^b
Midwest	14.44	15.90 ^b	18.73 ^b	14.23	15.09	18.73 ^b
South	13.86	15.33 ^b	17.44 ^b	13.45	14.64 ^a	17.44 ^b
West	13.41	17.23 ^b	17.94 ^b	13.18	16.49 ^b	17.94 ^b
Employment						
Full or Part Time	9.03	12.54 ^b	15.29 ^b	8.94	12.13 ^b	15.29 ^b
Unemployed	11.40	17.67 ^b	18.71 ^b	11.01	15.89 ^b	18.71 ^b
Other	21.35	20.57	21.66	20.84	19.75	21.66
Education						
< High School	21.57	21.10	18.46 ^b	20.71	20.03	18.46 ^b
High School Grad	15.67	17.53 ^b	18.04 ^b	15.43	16.68 ^a	18.04 ^b
Some College	12.85	16.24 ^b	19.38 ^b	12.53	15.21 ^b	19.38 ^b
College Grad	8.26	11.97 ^b	16.67 ^b	8.26	11.74 ^b	16.67 ^b
Poverty Level¹						
< 100% Threshold	25.72	27.44	24.67	24.83	25.90	24.67
100-199% Threshold	18.39	20.19 ^a	20.50 ^b	18.05	19.29	20.50 ^b
≥ 200% Threshold	10.47	12.63 ^b	16.01 ^b	10.30	12.10 ^b	16.01 ^b
Health Insurance						
Yes	12.96	15.05 ^b	17.58 ^b	12.73	14.37 ^b	17.58 ^b
No	16.98	21.55 ^b	20.42 ^b	16.46	20.46 ^b	20.42 ^b

AMI = any mental illness; Grad = graduate; K6 = Kessler-6; MDE = major depressive episode; MEPS = Medical Expenditure Panel Survey; SE = standard error; WHODAS = World Health Organization Disability Assessment Schedule.

^a Difference between this estimate and the estimate from the MEPS model in the column to the left is statistically significant from zero at the 0.05 level.

^b Difference between this estimate and the estimate from the MEPS model in the column to the left is statistically significant from zero at the 0.01 level.

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

Predictor variables included in the following models:

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

The predictor variables above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

Domains: MEPS had no equivalent variables to describe county type or receipt of mental health services, and there was ambiguity in how full-time employment and part-time employment were defined (so these categories were collapsed).

Datasets = 2010 adult NSDUH data, and 2010 adult MEPS public use data.

Analysis weight = ANALWT for NSDUH, and MEPS analysis weights.

Source: SAMHSA, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2010.

7. Summary and Conclusions

The objective of this National Survey on Drug Use and Health (NSDUH) methodological study was to investigate alternative serious mental illness (SMI) prediction models to see how they compare with the 2012 model to assess their usefulness in the following particular situation:

- One or more of the predictor variables in the SMI prediction model is also a key analytic variable with respect to an analysis of SMI.

Not all of the information required to construct the 2012 model is available (e.g., in other data sources). For this reason, a set of models with varying degrees of parsimony was investigated to see how the models would perform in terms of bias (overall and at the domain level), total error rate (sum of false-positive and false-negative rates), and how model-based estimates would compare against direct estimates as computed from the clinical sample. One set of six models includes the past year version of the Kessler-6 (K6) scale, and a second set of six models is identical except that the past month version of the K6 scale is collected instead (see Section 3.1 for detailed descriptions of the models).

Based on the various analyses discussed in this methodological document, the following broad conclusions can be drawn about the relationship between SMI and any mental illness (AMI) estimation and the models that were investigated.

SMI estimation:

- Models that do not include the World Health Organization Disability Assessment Schedule (WHODAS) term tended to exhibit higher bias and/or total error rate, suggesting that the inclusion of this term is important to control bias and the total error rate.
- The exclusion of suicidal thoughts or past year major depressive episode (MDE) from the model appeared to have little effect on the bias, but the total error rate showed some increase for each variable excluded.
- The use of past month K6 instead of past year K6 in the models did not seem to affect the bias or total error rate much, suggesting that past month K6 could be used without substantially affecting the predictions if the past year version were unavailable.

AMI estimation:

- All models that include past month K6 showed fairly large increases in bias, particularly those that do not include WHODAS, suggesting that the past month version of K6 in any model may lead to biased estimates of AMI.
- Models that include the past year K6 and WHODAS terms tended to exhibit reasonably low levels of bias and total error rate, and adding the age term (which was available in most surveys) improved those measures even further.

Therefore, the broad conclusions listed above suggest the following responses to the two objectives of this study:

- If predictor variables in the SMI prediction model (e.g., suicidal thoughts or past year MDE) are also key analytic variables with respect to a joint analysis with SMI or AMI, then analysts should consider using a model that does not include those predictor variables.
- For SMI and AMI prevalence estimates from other studies to be reasonably comparable with those from NSDUH, the other studies need to at least collect WHODAS information in addition to K6 information (with a preference for past year K6 information for AMI estimates). The absence of WHODAS information from the Behavioral Risk Factor Surveillance System (BRFSS), National Health Interview Survey (NHIS), and Medical Expenditure Panel Survey (MEPS) studies may explain in large part why SMI and AMI estimates obtained from those studies differ markedly from those obtained from NSDUH. However, even if any of those studies had collected WHODAS information, the NSDUH-based models may not be applicable to them because of the different methodological characteristics of the various studies.

In addition, a summary of the strengths and weaknesses of each of the 12 models with respect to SMI and AMI estimation is given in [Table 7.1](#).

Table 7.1 Strengths and Weakness of Models

Model	Strengths	Weaknesses
Y1	<ul style="list-style-type: none"> Requires only past year K6 in model and hence may be applicable to any source that collects information related to this variable, but not information related to other predictor variables, such as WHODAS, past year MDE, or suicidal thoughts. 	<ul style="list-style-type: none"> SMI and AMI estimates indicate a fair (but nonsignificant) degree of bias overall and a significant level of bias across a number of domains. Total error rate of SMI and AMI estimates is somewhat higher than it is for estimates based on the 2012 model.
Y2	<ul style="list-style-type: none"> Requires only past year K6 and age in model and hence may be applicable to as many sources as model Y1 because age is fairly universally collected. 	<ul style="list-style-type: none"> SMI and AMI estimates indicate a fair (but nonsignificant) degree of bias overall and a significant level of bias across a number of domains. Total error rate of SMI and AMI estimates is somewhat higher than it is for estimates based on the 2012 model, although it is also slightly smaller than it is for estimates based on model Y1.
Y3	<ul style="list-style-type: none"> Requires only past year K6 and WHODAS in model and hence may be applicable to any source that collects information related to these variables, but not information related to other predictor variables, such as past year MDE or suicidal thoughts. SMI estimate shows little bias overall, and AMI estimate shows reduced bias overall when compared with that of models Y1 and Y2. 	<ul style="list-style-type: none"> Requires terms in model that may not be collected in some sources, such as WHODAS. SMI estimates indicate a significant level of bias across different age groups; AMI estimates indicate a small (but nonsignificant) degree of bias overall and a significant level of bias across a small number of domains. Total error rate of SMI and AMI estimates is somewhat higher than it is for estimates based on the 2012 model, although it is also slightly smaller than it is for estimates based on models Y1 or Y2.
Y4	<ul style="list-style-type: none"> Requires only past year K6, WHODAS, and age in model and hence may be applicable to as many sources as model Y3 because age is fairly universally collected. SMI and AMI estimates show little bias overall. 	<ul style="list-style-type: none"> Requires terms in model that may not be collected in some sources, such as WHODAS. SMI estimates indicate a significant level of bias across one age group, and AMI estimates indicate a significant level of bias in one domain. Total error rate of SMI and AMI estimates is somewhat higher than it is for estimates based on the 2012 model, although it is also slightly smaller than it is for estimates based on models Y1 or Y2.
Y5	<ul style="list-style-type: none"> Requires past year K6, WHODAS, age, and past year MDE in model and hence may be applicable to any source that collects information related to these variables, but not information related to other predictor variables, such as suicidal thoughts. SMI and AMI estimates show little bias overall. SMI estimates indicate no significant level of bias at the domain level. Total error rate of SMI and AMI estimates is comparable or only slightly larger than that of 2012 model. Can jointly analyze mental illness and suicidality measures. 	<ul style="list-style-type: none"> Requires terms in model that may not be collected in some sources, such as WHODAS and past year MDE. AMI estimates indicate significant levels of bias for two domains.
Y6	<ul style="list-style-type: none"> SMI and AMI estimates show little bias overall. SMI estimates indicate no significant level of bias at the domain level, and AMI estimates indicate significant level of bias for only one domain. Total error rate of SMI and AMI estimates is smaller than for any other model. 	<ul style="list-style-type: none"> Requires terms in model that may not be collected in some sources, such as WHODAS, past year MDE, and suicidal thoughts. Should not jointly analyze mental illness and suicidality measures.

See notes at end of table.

(continued)

Table 7.1 Strengths and Weakness of Models (continued)

Model	Strengths	Weaknesses
M1	<ul style="list-style-type: none"> Similar to model Y1, except may be applicable to more sources (i.e., to those that collect past month K6 instead of past year K6). 	<ul style="list-style-type: none"> Similar to model Y1 with respect to SMI. Bias substantially higher for AMI estimates when compared with model Y1.
M2	<ul style="list-style-type: none"> Similar to model Y2, except may be applicable to more sources (i.e., to those that collect past month K6 instead of past year K6). 	<ul style="list-style-type: none"> Similar to model Y2 with respect to SMI. Bias substantially higher for AMI estimates when compared with model Y2.
M3	<ul style="list-style-type: none"> Similar to model Y3, except may be applicable to more sources (i.e., to those that collect past month K6 instead of past year K6). 	<ul style="list-style-type: none"> Similar to model Y3 with respect to SMI. Bias substantially higher for AMI estimates when compared with model Y3.
M4	<ul style="list-style-type: none"> Similar to model Y4, except may be applicable to more sources (i.e., to those that collect past month K6 instead of past year K6). 	<ul style="list-style-type: none"> Similar to model Y4 with respect to SMI. Bias substantially higher for AMI estimates when compared with model Y4.
M5	<ul style="list-style-type: none"> Similar to model Y5, except may be applicable to more sources (i.e., to those that collect past month K6 instead of past year K6). 	<ul style="list-style-type: none"> Similar to model Y5 with respect to SMI. Bias substantially higher for AMI estimates when compared with model Y5.
M6	<ul style="list-style-type: none"> Similar to model Y6, except may be applicable to more sources (i.e., to those that collect past month K6 instead of past year K6). 	<ul style="list-style-type: none"> Similar to model Y6 with respect to SMI. Bias substantially higher for AMI estimates when compared with model Y6.

AMI = any mental illness; K6 = Kessler-6; SMI = serious mental illness; MDE = major depressive episode; SMI = serious mental illness; WHODAS = World Health Organization Disability Assessment Schedule.

Predictor variables included in the following models:

Y1: K6 (Past Year)

Y2: K6 (Past Year) + Age

Y3: K6 (Past Year) + WHODAS

Y4: K6 (Past Year) + WHODAS + Age

Y5: K6 (Past Year) + WHODAS + Age + Past Year MDE

Y6: K6 (Past Year) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

M1: K6 (Past Month)

M2: K6 (Past Month) + Age

M3: K6 (Past Month) + WHODAS

M4: K6 (Past Month) + WHODAS + Age

M5: K6 (Past Month) + WHODAS + Age + Past Year MDE

M6: K6 (Past Month) + WHODAS + Age + Past Year MDE + Suicidal Thoughts

The predictor variables listed above are defined in exactly the same way as in the 2012 model (see Chapter 2): K6 (Past Year), WHODAS, Age, Suicidal Thoughts, and Past Year MDE. The K6 (Past Month) variable is defined analogously to the K6 (Past Year) variable; the only difference is that past month data are used instead. The model Y6 (in bold) describes the 2012 model.

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

References

- Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. (n.d.). *Medical Expenditure Panel Survey*. Retrieved from <http://meps.ahrq.gov/mepsweb/>
- Center for Behavioral Health Statistics and Quality. (2009). *2008 National Survey on Drug Use and Health: Methodological resource book (Section 16, Measuring serious mental illness with the NSDUH: Results of 2008 12-month analysis)*. 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)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Center for Behavioral Health Statistics and Quality. (2015). *2014 National Survey on Drug Use and Health: Methodological summary and definitions*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Centers for Disease Control and Prevention. (2014, December 4). *BRFSS 2012 survey data and documentation*. Retrieved from http://www.cdc.gov/brfss/annual_data/annual_2012.html
- Centers for Disease Control and Prevention. (2015, September 1). *National Health Interview Survey*. Retrieved from <http://www.cdc.gov/nchs/nhis.htm>
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. L., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*, 184-189. doi:10.1001/archpsyc.60.2.184
- 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. doi:10.2190/T8EM-C8YH-373N-1UWD
- National Center for Health Statistics, Division of Health Interview Statistics. (2013, June). *2012 National Health Interview Survey (NHIS) public use data release: NHIS survey description*. Retrieved from ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2012/srvydesc.pdf
- 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

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