This report was prepared by the Substance Abuse and Mental Health Services Administration using federal funding provided by the Bureau of Justice Statistics.

Document Title:	Statistical Models to Predict Mental Illness Among State and Federal Prisoners
Authors:	Jeremy Aldworth, RTI International Phillip Kott, RTI International Sarra Hedden, formerly Substance Abuse and Mental Health Services Administration Joseph Gfroerer, former consultant, Bureau of Justice Statistics Lauren G. Beatty, Bureau of Justice Statistics
Document No.:	252633
Publication Date:	June 2021
Award No.:	This project was supported by award number HHSS283201000003C

Abstract:

The Substance Abuse and Mental Health Services Administration (SAMHSA) developed a model-based methodology to obtain nationally representative estimates of serious mental illness (SMI) and any mental illness (AMI) among the adult U.S. civilian, non-institutionalized population in the 2008-12 Mental Health Surveillance Study (MHSS) as part of the National Survey on Drug Use and Health (NSDUH). This paper examines the feasibility of adapting and applying this methodology to 2016 Survey of Prison Inmates (SPI) data, collected by the Bureau of Justice Statistics (BJS), to obtain nationally representative estimates of SMI and AMI among state and federal prisoners.

The general methodology maintains a rigorous approach that provides reasonably accurate estimates of SMI and AMI (both overall and at the domain level). However, estimates obtained from the 2016 SPI should be viewed with some caution (particularly at the domain level, e.g., sex or age group). Several potential limitations of the methodology when specifically applied to the SPI arise from the small sample size of the MHSS subsample used to develop the SPI prediction model, limited information in the 2016 SPI available for inclusion in prediction models, and different populations represented in the model development process (parolees, probationers, and arrestees surveyed in NSDUH) and estimation phase (federal and state prisoners surveyed in SPI).

Despite these limitations, this approach to estimating mental illness among prisoners may provide useful indicators that are consistent with estimates produced by other federal agencies and researchers. Estimates of SMI and AMI are higher among the prisoner population than among the general adult population not in prison, large differences exist between federal and state prisoners and between male and female prisoners, and minimal differences exist by age group.

Disclaimer

This page intentionally left blank.

Statistical Models to Predict Mental Illness Among State and Federal Prisoners

Jeremy Aldworth and Phillip Kott
RTI International

Sarra Hedden
Formerly with the Substance Abuse and Mental Health Services Administration

Joseph Gfroerer
Former Consultant for the Bureau of Justice Statistics

Lauren G. Beatty
Bureau of Justice Statistics

NCJ 252633, June 2021

Abstract (or Highlights)

The Substance Abuse and Mental Health Services Administration (SAMHSA) developed a model-based methodology to obtain nationally representative estimates of serious mental illness (SMI) and any mental illness (AMI) among the adult U.S. civilian, non-institutionalized population in the 2008-12 Mental Health Surveillance Study (MHSS) as part of the National Survey on Drug Use and Health (NSDUH). This paper examines the feasibility of adapting and applying this methodology to 2016 Survey of Prison Inmates (SPI) data, collected by the Bureau of Justice Statistics (BJS), to obtain nationally representative estimates of SMI and AMI among state and federal prisoners.

The general methodology maintains a rigorous approach that provides reasonably accurate estimates of SMI and AMI (both overall and at the domain level). However, estimates obtained from the 2016 SPI should be viewed with some caution (particularly at the domain level, e.g., sex or age group). Several potential limitations of the methodology when specifically applied to the SPI arise from the small sample size of the MHSS subsample used to develop the SPI prediction model, limited information in the 2016 SPI available for inclusion in prediction models, and different populations represented in the model development process (parolees, probationers, and arrestees surveyed in NSDUH) and estimation phase (federal and state prisoners surveyed in SPI).

Despite these limitations, this approach to estimating mental illness among prisoners may provide useful indicators that are consistent with estimates produced by other federal agencies and researchers. Estimates of SMI and AMI are higher among the prisoner population than among the general adult population not in prison, large differences exist between federal and state prisoners and between male and female prisoners, and minimal differences exist by age group.

Glossaries

General abbreviations

Abbreviation	Description
ACASI	Audio computer-assisted self-interviewing
AMI	Any mental illness
AUC	Area under ROC curve
BJS	Bureau of Justice Statistics
CAPI	Computer-assisted personal interviewing
CBHSQ	Center for Behavioral Health Statistics and Quality
CDF	Cumulative distribution function
CMHS	Center for Mental Health Services
CJ	Criminal justice
DF	Degrees of freedom
DSM	Diagnostic and Statistical Manual of Mental Disorders
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, fourth edition
ECA	Epidemiologic Catchment Area
FN	False negative
FP	False positive
GAF	Global Assessment of Functioning
K6	Kessler 6-item distress scale
ICD-9-CM	International Classification of Diseases, Ninth Revision, Clinical Modification
LMI	Mild (or low) mental illness
MDE	Major depressive episode
MEPS	National Medical Expenditures Panel Survey
MHSS	Mental Health Surveillance Study
MMI	Moderate mental illness
NCS	National Comorbidity Study
NHIS	National Health Interview Survey
NIMH	National Institute of Mental Health
NIS	National Inmate Survey
NPV	Negative predictive value
NSDUH	National Survey on Drug Use and Health
PAPI	Paper-and-pencil interviewing
PP	Parolees or probationers
PPV	Positive predictive value
PSU	Primary sampling unit
ROC	Receiver operating characteristic
SAMHSA	Substance Abuse and Mental Health Services Administration
SCID	Structured Clinical Interview for DSM-IV
SDS	Sheehan Disability Scale
SE	Standard error
SMI	Serious mental illness
SPD	Serious psychological distress
SPI	Survey of Prison Inmates
TAG	Technical advisory group
TCE	Total classification error
WHODAS	World Health Organization Disability Assessment Schedule

Variable names and weights (variables specific to SPI are italicized)

Name	Description
AGE18	Coded as age minus 18
AGE1830	Coded as age minus 18 if age 18 to 30; coded as 12 otherwise
AGE1845	Coded as age minus 18 if age 18 to 45; coded as 27 otherwise
AGE1850	Coded as age minus 18 if age 18 to 50; coded as 32 otherwise
AGE20	Coded as 0 for all ages up to age 20, then as age minus 20 thereafter
AGE25	Coded as 0 for all ages up to age 25, then as age minus 25 thereafter
AGE2545	Coded as 0 for all ages up to age 25; coded as age minus 25 if age 25 to 45; coded as 20 otherwise
AGE30	Coded as 0 for all ages up to age 30, then as age minus 30 thereafter
AGE3040	Coded as 0 for all ages up to age 30; coded as age minus 30 if age 30 to 40; coded as 10 otherwise
AGE3045	Coded as 0 for all ages up to age 30; coded as age minus 30 if age 30 to 45; coded as 15 otherwise
AGE3050	Coded as 0 for all ages up to age 30; coded as age minus 30 if age 30 to 50; coded as 20 otherwise
AMDEY_2	Past year MDE (coded as 1 if “yes”; coded as 0 otherwise)
AMIYR_U	Coded as 1 if SMIPP_U greater than or equal to NSDUH AMI cut point; coded as 0 otherwise
ANALWT	Analysis weight for 2008-2012 adult NSDUH data
ANALWT_A	Analysis weight for 2008A-2012 adult NSDUH data
ANXDSLIF_U	Reported having had anxiety in lifetime (coded as 1 if “yes”; coded as 0 otherwise [i.e., includes “no” and missing values])
ANXLIF	*Reported ever having an anxiety disorder [SPI Item: MH7e] (coded as 1 if “yes”; coded as 0 if “no”)*
DEPLIF	*Reported ever having a depressive episode [SPI Item: MH7b] (coded as 1 if “yes”; coded as 0 if “no”)*
DEPRSLIF_U	Reported having depression in lifetime (coded as 1 if “yes”; coded as 0 otherwise [i.e., includes “no” and missing values])
K6SCMON	Past month K6 score
K6SCMON2	Alternative past month K6 score (K6 score of less than 8 recoded as 0; K6 score of 8 to 24 recoded as 1 to 17)
MHFAAWGT	Analysis weight for 2008A-2012 MHSS clinical data
MHFNLWGT	Analysis weight for 2008-2012 MHSS clinical data
MHSUTK_K	Serious thoughts of suicide during the past year (coded as 1 if “yes”; coded as 0 otherwise)
SCID_AMI	SCID (i.e., gold standard) determination of AMI (yes or no)
SCID_SMI	SCID (i.e., gold standard) determination of SMI (yes or no)
SMIPP_U	Predicted probability of SMI, based on 2012 NSDUH prediction model
SMIYR_U	Coded as 1 if SMIPP_U greater than or equal to NSDUH SMI cut point; coded as 0 otherwise
WHODASC3	WHODAS item score of less than 2 recoded as 0; WHODAS item score of 2 to 3 recoded as 1, then summed for a score ranging from 0 to 8
WSPDSC2	Alternative past year K6 score (K6 score of less than 8 recoded as 0; K6 score of 8 to 24 recoded as 1 to 17)
WT_FINAL	*Analysis weight for 2016 SPI data*

Models (models specific to SPI are italicized)

Name	**Predictor Variables in Model**
2012 NSDUH prediction model	WSPDSC2, WHODASC3, MHSUTK_U, AMDEY2_U, AGE1830
K2	K6SCMON2
K2A	K6SCMON2, ANXDLIF_U
K2A18	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE18
K2A20	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE20
K2A25	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE25
K2A30	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE30
K2A1845	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE1845
K2A1850	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE1850
K2A2545	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE2545
K2A3040	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE3040
K2A3045	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE3045
K2A3050	K6SCMON2, DEPRSLIF_U, ANXDLIF_U, AGE3050
K2D	K6SCMON2, DEPRSLIF_U
K2DA	K6SCMON2, DEPRSLIF_U, ANXDLIF_U
K2DA	*K6SCMON2, DEPLIF, ANXLIF*

Table of Contents

List of Figures

Figure **Page**

List of Tables

Table **Page**

1. Introduction

1.1 Background

The Bureau of Justice Statistics (BJS) provides the only source of national estimates on mental illness and mental health treatment needs of state and federal prisoners in the United States. BJS produces reliable estimates to assess the level and patterns of mental illness among state and federal prisoners, monitor trends, identify potential treatment and service needs, and inform policy and criminal justice planning. Historically, BJS has relied primarily on its national omnibus survey of state and federal prisoners, renamed as the Survey of Prison Inmates (SPI) with the 2016 iteration, as the primary data source to generate those estimates.[1]

However, methodologies used by BJS to measure the prevalence of mental illness among the prison populations through the SPI have changed over time, as BJS has tried to improve measurement. The use of different methodologies limits comparisons over time and has resulted in some estimates that were of limited value. (See Section 1.3 for a discussion on BJS methods to measure mental illness among prisoners over time.) For example, in the 2004 Survey of Inmates in State and Federal Correctional Facilities (SISFCF), the scope of mental health problems was broad and included prisoners with a recent history of a disorder or symptoms that were consistent with a mental disorder that would require further evaluation. Other federal population surveys focus on more narrowly defined and commonly used measures such as serious mental illness (SMI) (which requires both the presence of a disorder and serious impairment due to the disorder) or serious psychological distress (SPD). For instance, the National Survey on Drug Use and Health (NSDUH), administered by the Substance Abuse and Mental Health Services Administration (SAMHSA), provides the primary source of national and state-level data on mental health among the U.S. civilian, non-institutionalized population. In a correctional setting, persons with these more severe conditions constitute the higher priority groups because they are most at risk to themselves, other prisoners, and correctional staff. They are also most in need of the limited resources and services available through correctional agencies.

In recent years, BJS has sought to enhance the measurement of mental illness among the prison populations through a variety of ways, including the redesign of SPI and methodological research. These efforts have aimed to improve the validity of BJS estimates, comport with current practices and methods to measure this topic among the general population, and enhance the utility of its statistics to stakeholders. BJS has sought to adopt approaches to measure mental illness that are consistent with other federal population surveys and methods, such as SAMHSA's NSDUH. In the 2016 SPI, BJS included the past month K6 distress scale (Kessler et al., 2003), which is widely used in mental health epidemiologic studies, including the NSDUH, National Health Interview Survey (NHIS), and National Medical Expenditures Panel Survey (MEPS). Thus, BJS can now produce direct estimates of SPD (i.e., defined by a K6 score of 13 or more) among the state and federal prison populations.

[1] Prior iterations of the SPI were known as the Survey of Inmates in State and Federal Correctional Facilities. In recent years, BJS has also relied on the National Inmate Survey, a different national survey of prisoners and jail inmates, to produce estimates of mental illness among prisoners during the years that the SPI was not fielded.

While SAMHSA produces estimates of SPD among the general population, these estimates are not SAMHSA's primary indicator of mental health status. For mental health block grant allocation, policy, and research purposes, and as required by law, SAMHSA relies on estimates of SMI. This term is defined as any mental disorder (excluding substance use disorder) that results in serious functional impairment, according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria. Prior to 2005, SAMHSA used a score of 13 or more on the past year K6 scale as the criterion for estimating SMI. The cut point of 13 or more on the K6 is based on a predictive model developed from a small clinical sample (Kessler et al., 2003). However, in 2005, SAMHSA discontinued reporting estimates of SMI due to concerns about the accuracy of that model, in part because it lacked an impairment measure. SAMHSA continued to produce estimates of the prevalence of adults scoring 13 or more on the K6 scale, but these estimates were referred to as SPD. Later, SAMHSA conducted methodological research to construct an impairment scale to include in the NSDUH and develop an improved predictive model to estimate the prevalence of SMI among the U.S. civilian, non-institutionalized population age 18 or older (Center for Behavioral Health Statistics and Quality [CBHSQ], 2014a). The model also was used to produce estimates of any mental illness (AMI). Estimates of SMI and AMI based on this model-based method have been produced by SAMSHA since 2008.[2]

Starting in 2017, BJS collaborated with SAMHSA and its NSDUH contractor, RTI International, to explore the feasibility of adopting the NSDUH methodology to develop and evaluate a predictive model to estimate SMI and AMI among state and federal prisoners. These represent the two principal mental health indicators that SAMHSA relies on to describe the level and scope of mental illness in the United States from the NSDUH. This collaboration with SAMHSA and RTI International was an extension of BJS's efforts to improve its measurement of mental illness among prisoners and enhance the utility and relevance of its estimates on this topic to key stakeholders, including policymakers, mental health and corrections administrators, practitioners, and researchers.

The NSDUH predictive model used to produce estimates of SMI and AMI among the general adult population could not be applied to the 2016 SPI data to produce estimates of SMI among the prisoner populations. First, the 2016 SPI did not collect all the information required to create all the predictor variables in the NSDUH predictive model. For example, the NSDUH predictor model contains a measure of past-year depression while the SPI dataset does not (see Sections 2.7 and 3.4 for details). Preliminary analyses (not shown) indicated that coefficients of predictor variables created from measures of psychological distress, depression, and anxiety differed when applied to the general population versus the criminal justice population (i.e., probationers, parolees, and arrestees; see Section 4.1.1 for details). The relationship between predictor variables and SMI or AMI may differ between the general and prison populations, indicating that a new model should be fit to the prison populations.

Although the 2016 SPI questionnaire does not include an impairment scale due to constraints of the length of the interview, there are several items in the SPI that appeared to be reasonable candidates for inclusion in predictive models assessed under this project. The SMI and AMI estimates of state and federal prisoners presented in this report will be labeled as developmental until further review and verification by BJS to fully endorse the methodology. Although the estimates in this report are

[2] Definitions for SMI and AMI are given in Section 1.4.

labeled as “developmental,” the methodology behind them has gone through rigorous review and verification.

1.2 Purpose of the Report

The purpose of this report is to—

- develop and evaluate a methodology for estimating SMI and AMI among the state and federal prison populations
- provide a detailed description of the methodology.

1.3 Prior BJS Research on Mental Illness Among Prisoners

From 1999 to 2017, BJS published three reports on mental illness among prisoners. In each of these reports, the definition of mental illness and data source differed, thus the prevalence estimates of mental illness among state and federal prisoners also varied.

The first report (Ditton, 1999) analyzes mental illness among prisoners based on the 1997 SISFCF. Prisoners were identified as mentally ill if they reported one of the following criteria: (1) current mental or emotional condition; or (2) overnight stay in a mental hospital, unit, or treatment program during their lifetime. The prevalence of mental illness among state prisoners was estimated at 16% and 7% for federal prisoners.

The second report (James & Glaze, 2006) examines mental illness among prisoners based on 2004 SISFCF data. Mental health problems were defined by (1) a recent history of mental health problems during the year before their arrest or since admission to prison or (2) symptoms of a mental disorder that occurred during the 12 months prior to the interview. A recent history of mental health problems included having ever been told by a mental health professional as having a mental disorder or receiving treatment, including being prescribed medication; staying overnight in a mental hospital, unit, or treatment program; or receiving other treatment from a mental health professional. Symptoms of a mental disorder were based on criteria specified in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV). The prevalence of mental illness among state prisoners was estimated at 56% and 45% for federal prisoners.

The third report (Bronson & Berzofsky, 2017) evaluates mental illness among prisoners based on the 2011-12 National Inmate Survey (NIS-3). Two mental illness measures were presented: (1) SPD during the past 30 days as measured by the K6 nonspecific psychological distress scale (Kessler et al., 2003); and (2) a history of a mental health problem, which was defined as having ever been told by a mental health professional as having a mental disorder. State and federal prisoner prevalence estimates were combined into a single estimate for each of the two mental illness measures: 14% with SPD and 37% with a history of a mental health problem.

1.4 Brief Description of Approach Used in this Report

To obtain estimates of SMI and AMI among the prisoner populations, this report follows the approach used to obtain the same estimates among the general population by CBHSQ (2014a). CBHSQ (2014a) describes how, in 2008, SAMHSA implemented a new program to produce national and state estimates of the prevalence of mental illness. The Mental Health Surveillance Study (MHSS) was conducted as part of the NSDUH from 2008 to 2012 for the primary purpose of developing models that estimate the prevalence of SMI and AMI in the U.S. adult (18 or older) civilian, non-institutionalized population. The MHSS consisted of three major components:

- a new questionnaire module administered to all adult respondents that captured data on psychological distress, impairment caused by mental health issues, and suicidality (thoughts, plans, and attempts)
- a subsample of adults selected from the main NSDUH study for follow-up clinical interviews
- development of a prediction model and cut point in the MHSS subsample that was applied to the NSDUH sample, i.e., variables collected in the main NSDUH survey were used in the prediction model to generate estimates of mental illness.

Information from the clinical interviews was used to determine the gold-standard measure of SMI status ("yes" or "no") for each subsample respondent (see Section 2.4 for details). The clinical interview data were linked with the data obtained from the main NSDUH survey to develop a prediction model to estimate SMI and AMI for the full adult NSDUH sample. A prediction model was developed that included gold-standard SMI status (obtained from the clinical interviews) as the response variable. The predictor variables included variables such as psychological distress, impairment caused by mental health issues, major depressive episode (MDE), thoughts of suicide, and age that were collected in the main NSDUH questionnaire for adults. Once a model was determined, all respondents in the MHSS clinical sample had both a predicted probability of SMI *and* gold-standard measures of SMI and AMI status. Next, a cut point was determined (in the clinical sample). 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. 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 approximately equal, thus ensuring unbiased estimates. Using the same model, a cut point for producing unbiased estimates of AMI was also determined. Finally, the SMI prediction model and cut point determined from the MHSS clinical data was then extrapolated to all adult respondents in the main NSDUH sample to obtain the predicted probability of SMI, and predicted SMI and AMI status, for each of those respondents. See Chapter 2 for more details.

For the purpose of obtaining estimates of SMI and AMI among the prison populations, modifications to the modeling and estimation approach used in the MHSS were necessary because the state and federal prison populations are omitted from the NSDUH, and the SPI did not include the same predictor variables in the NSDUH model. The modifications to the approach included the following:

- Criminal justice subpopulations covered in the NSDUH to serve as a proxy for the prison populations were identified, and analyses were restricted to these subpopulations.

The criminal justice subpopulations identified in the NSDUH included parolees, probationers, and arrestees.[3]

- Because the models developed within the NSDUH would ultimately be extrapolated to prisoner data contained within the 2016 SPI, it was necessary to use predictor variables common to both the NSDUH and SPI.

1.5 Organization of this Report

Chapter 2 discusses the NSDUH method of modeling and estimating SMI and AMI through the MHSS. Chapter 3 describes the purpose, history, design, response rates, and mental health items of the 2016 SPI. Chapter 4 details the approach used to model and estimate SMI, and Chapter 5 reviews the modeling and estimation of AMI. Chapter 6 examines the limitations of the approach used in this report and guidelines for use. Finally, Chapter 7 provides a summary and concluding remarks.

[3] The specific questions included in the NSDUH to identify these subpopulations were: (1) Were you on parole, supervised release, or other conditional release from prison at any time during the past 12 months?; (2) Were you on probation at any time during the past 12 months?; (3a) Not counting minor traffic violations, have you ever been arrested and booked for breaking the law?; and (3b) Not counting minor traffic violations, how many times during the past 12 months have you been arrested and booked for breaking a law?

2. NSDUH Method of Estimating SMI and AMI

2.1 Summary of NSDUH and Need for Mental Health Data

NSDUH annually surveys the U.S. civilian, non-institutionalized population age 12 or older and collects data on the use of illegal drugs, alcohol, tobacco, and mental health. Since 2008, the mental health modules included a psychological distress scale, impairment scale, module to assess past year and lifetime MDE, questions on suicidality (thoughts, plans, and attempts), modules on experiences with treatment for mental health, and, until 2014, questions on whether respondents had been told by a medical professional that they had depression or anxiety disorder. Conducted by the federal government since 1971, the survey collects data through face-to-face interviews[4] with a representative sample of the population at the respondent's place of residence. The survey is sponsored by SAMHSA, U.S. Department of Health and Human Services, and is planned and managed by SAMHSA's CBHSQ. Data collection and analysis are currently conducted under contract with RTI International.[5]

NSDUH collects information from residents of households and non-institutional group quarters (e.g., shelters, rooming houses, dormitories) and from civilians living on military bases. The survey *excludes* homeless persons who do not use shelters, military personnel on active duty, and long-term residents of institutional group quarters, such as prisons, jails, and hospitals.

The NSDUH provides annual estimates of SMI for the non-institutionalized population age 18 or older. As described in Chapter 1, the MHSS was conducted as part of the NSDUH from 2008 to 2012 for the primary purpose of developing models that estimate the prevalence of SMI and AMI in adults at the national and state levels. The Alcohol, Drug Abuse, and Mental Health Administration Reorganization Act of 1992 (P.L. 102-321) established a block grant for U.S. states to fund community mental health services for adults with SMI. The act required states to include prevalence estimates in their annual applications for block grant funds. This legislation also mandated SAMHSA to develop an operational definition of SMI and a methodology to produce national and state estimates. The MHSS clinical follow-up study was conducted to develop a model to generate estimates of SMI.

SAMHSA (2009), which administers the Center for Mental Health Services (CMHS), published its definition of SMI in 1993 in the *Federal Register*:

> Pursuant to Section 1912(c) of the Public Health Services Act, as amended by Public Law 102-321, "adults with serious mental illness" are defined as the following:

---

[4] Each NSDUH interview includes both computer-assisted personal interviewing (CAPI), in which interviewers administer the questions to respondents, and audio computer-assisted self-interviewing (ACASI), in which respondents answer the questions without having to reveal their answers to interviewers. ACASI is designed to maximize accurate reporting of sensitive information about substance use and mental health by providing respondents with a highly private and confidential mode for responding to questions. Less sensitive items are administered by interviewers using CAPI.
[5] RTI International is a registered trademark and a trade name of Research Triangle Institute.

- Persons aged 18 and over, who currently or at any time during the past year, have had diagnosable mental, behavioral, or emotional disorders of sufficient duration to meet diagnostic criteria specified within DSM-III-R [sic] that has resulted in functional impairment, which substantially interferes with or limits one or more major life activities.
- These disorders include any mental disorders (including those of biological etiology) listed in DSM-III-R or their International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) equivalent (and subsequent revisions), with the exception of DSM-III-R “V” codes, substance use disorders, and developmental disorders, which are excluded unless they co-occur with other diagnosable serious mental illness.
- All of these disorders have episodic, recurrent, or persistent features; however, they vary in terms of severity or disabling effects. Functional impairment is defined as difficulties that substantially interfere with or limit role functioning in one or more major life activities, including basic daily living skills (e.g., eating, bathing, dressing); instrumental living skills (e.g., maintaining a household, managing money, getting around the community, taking prescribed medication); and functioning in social, family, and vocational/educational contexts.
- Adults who would have met functional impairment criteria during the referenced year without benefit of treatment or other support services are considered to have serious mental illnesses. (p. 147)

Following the publication of the SMI definition, SAMHSA struggled with finding a valid methodology. Data on mental illness in the population are scarce, and measuring comparable and nationally representative data on mental illness at a detailed level and on an annual basis at the national and state level present challenges. Initially, in 1990, SAMHSA developed a model-based estimation approach using data from the National Comorbidity Study (NCS) and the Epidemiologic Catchment Area (ECA) study, which was conducted in the early 1980s and not nationally representative (Hedden et al., 2012). In 2001, SAMHSA added the K6 scale and other mental health indicators to the NSDUH, and it conducted a small (n < 100) clinical interview follow-up study in one local area to develop a predictive model for SMI (Kessler et al., 2003). Annual estimates of SMI prevalence based on this model were published for 3 years. However, in 2004, SAMHSA discontinued the approach due to concerns about its validity, in part due to the lack of an impairment indicator in the predictive model. The predictive model was a single-variable model, which simply classified a respondent as having SMI if his or her past year K6 score was 13 or more. SAMHSA continued to publish prevalence estimates for those with K6 score of more than 12, but they have since been referred to as SPD.

In December 2006, a technical advisory group (TAG) meeting of expert consultants was convened by SAMHSA to solicit recommendations for mental health surveillance data collection strategies among the U.S. population. The panel recommended that the NSDUH be used to develop a better methodology for producing estimates of SMI among adults by including short scales in the main interview that are strong predictors of SMI (including an impairment indicator). They also suggested administering a gold-standard clinical psychiatric interview to a large, nationally representative subset of respondents to provide the data for estimating a statistical model that predicts SMI. In response, SAMHSA's CBHSQ initiated the MHSS as part of the NSDUH to develop and implement a valid method to estimate SMI in 2008.

2.2 Implementation of the MHSS (2008-2012)

The implementation of the MHSS required that strong predictors of SMI be available from all adults in the NSDUH. These predictors of SMI would include proxies for the presence of a relevant mental disorder and functional impairment due to having a mental health issue.

In consultation with the TAG, two candidate impairment scales were selected by SAMHSA to be added to the 2008 NSDUH: (1) an abbreviated version of the World Health Organization Disability Assessment Schedule (WHODAS; Rehm et al., 1999), and (2) the Sheehan Disability Scale (SDS; Leon et al., 1997). Initially, these scales were modified for use in a general population survey, including changes to question wording and length, which resulted in an abbreviated eight-item version of the WHODAS (Novak et al., 2010). To determine which impairment scale better predicted SMI, a split-sample design was used in the 2008 NSDUH. All adult respondents to the main NSDUH interview received the K6, a random half of the sample received the WHODAS (i.e., Sample A), and the other half received the SDS (i.e., Sample B).

The MHSS also included a diagnostic clinical interview administered to a subsample of adult NSDUH respondents. In 2008 (the first year of the MHSS), a subsample of approximately 1,500 adult NSDUH participants completed a follow-up clinical interview. The randomization of the impairment scales was maintained within this clinical interview subsample, which this report refers to as the clinical sample. About half of the MHSS clinical sample participants were administered the WHODAS, and half were administered the SDS (i.e., there were approximately 750 completed interviews from each half sample). Each participant in the 2008 MHSS clinical study was administered the Structured Clinical Interview for DSM-IV (SCID)[6] over the telephone with the clinical interviewers using paper-and-pencil interviewing (for further details of the SCID, see Appendix H of CBHSQ (2014b)). The clinical interviews were administered approximately 2 to 4 weeks after the main NSDUH interview. Functional impairment ratings were assigned by clinical interviewers using the Global Assessment of Functioning (GAF) scale[7]. The model estimation analyses used the clinical SCID/GAF combination as the indicator of SMI in evaluating which combination of K6 and impairment scale worked best in the statistical model used to predict SMI status.

Based on an analysis of the 2008 MHSS data, it was determined that the WHODAS would be administered as the sole impairment scale in subsequent NSDUHs (starting in 2009); see Aldworth

[6] The Structured Clinical Interview for the DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP) (First, Spitzer, Gibbon, & Williams, 2002).
[7] The GAF is a numeric scale (0 through 100) used to subjectively rate the social, occupational, and psychological functioning of adults. The DSM-IV-TR presents and describes the GAF (see p. 32 of American Psychiatric Publishing, Inc., 2000; also see Endicott, Spitzer, Fleiss, & Cohen, 1976). Lower scores represent higher levels of functional impairment. Descriptions of impairment are provided at 10-point intervals (e.g., 1 to 10, 11 to 20, and so on up to 91 to 100). For example, a GAF score between 51 and 60 is described as having moderate symptoms of impairment, while a score higher than 60 represents several categories of impairment ranging from none to slight. A score lower than 51 represents several categories ranging from serious to extreme.

et al. (2009) for more details. The MHSS was implemented over 5 years (2008-2012). The 2009 and 2010 samples were designed to yield 500 clinical interviews, and the 2011 and 2012 samples were designed to yield 1,500 clinical interviews, giving a 5-year total of 5,500 clinical interviews. By the end of 2012, a nationally representative sample of 5,653 completed clinical assessments included the 1,500 respondents from 2008, 520 from 2009, 516 from 2010, 1,495 from 2011, and 1,622 from 2012.

The subset of respondents to whom the WHODAS questions were assigned and who provided completed clinical assessments included 759 respondents from 2008, and all other respondents with completed assessments from 2009 to 2012. Therefore, from 2008 to 2012 there was a nationally representative sample of 4,912 completed clinical assessments for respondents assigned to the WHODAS.

Details of the three scales (i.e., K6, WHODAS, and SDS) added to the main interview of the 2008 NSDUH are given in Section 2.3, and details of the SCID are provided in Section 2.4.

2.3 NSDUH Mental Health Scales Used to Create Predictors of SMI

2.3.1 Psychological Distress Scale (K6)

Used to capture nonspecific psychological distress (Kessler et al., 2004), the K6 scale consists of two sets of six questions that ask respondents how frequently they experienced symptoms of psychological distress during two different time periods: (1) during the past 30 days and (2) the one month during the past year when they were at their worst emotionally. Respondents were asked about the second time period if they indicated that there was one month during the past 12 months when they felt more depressed, anxious, or emotionally stressed than they felt during the past 30 days. The six domains covered by the questions corresponded to how often the respondent felt (1) nervous, (2) hopeless, (3) restless or fidgety, (4) sad or depressed, (5) that everything was an effort, and (6) worthless. To create a score, the six items related to the first time period were coded from 0 to 4: “all of the time” was coded as 4; “most of the time” as 3; “some of the time” as 2; “a little of the time” as 1; and “none of the time,” “don't know,” and “refuse” coded as 0. Summing across the six responses resulted in a total score of 0 to 24. The six items related to the second time period were coded identically. The worst K6 total score was calculated as the maximum of the total scores from the two time periods and is considered the past year K6 total score. An alternative version of the past year K6 total score was formulated as follows: past year K6 total scores of less than 8 were recoded as 0, and past year K6 total scores from 8 to 24 were recoded as 1 to 17. The alternative version was used because SMI prevalence was typically extremely low for respondents with worst K6 total scores less than 8, and the prevalence rates started increasing once total scores were 8 or more. Therefore, a score band of 0-17 collapsed the less informative lower scores into a one-score category (0) while preserving the more informative scores at the higher end of the scale (1-17) (Aldworth et al., 2010).

2.3.2 Functional Impairment Scales (WHODAS and SDS)

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

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

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

One exception to this coding related to the last WHODAS item on how much difficulty respondents had in getting their daily work done as quickly as needed. If they responded to the fifth category of the *previous question* (i.e., that they did not go to work or school), then they skipped over the last item. Their response to this item was recorded as their emotions, nerves, or mental health caused them to be unable to go to work or school. Otherwise, this item was coded similarly to the other items.

Summing across the eight responses resulted in a total score of 0 to 24. An alternative version of the WHODAS total score was formulated as follows: item scores of less than 2 were recoded as 0, and item scores from 2 to 3 were recoded as 1, then summed for a total score ranging from 0 to 8. An alternative version of the WHODAS total score that uses a dichotomous measure dividing respondents who experienced moderate or severe difficulties from the remaining respondents might better predict SMI than a linear continuous measure.

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

2.4 Clinical Follow-Up Instruments

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

Table 1 lists diagnostic modules in the MHSS version of the SCID. The assessment of lifetime manic episode was included to provide a context for understanding whether a person experienced a past year major depressive episode (MDE) as part of a unipolar mood disorder or as a component of a bipolar disorder (regardless of whether a manic episode also was experienced during the past year). The module to assess intermittent explosive disorder was obtained from the optional impulse control disorders section of the SCID. Although the module for substance use disorders was administered to respondents, substance use disorder is not included in SAMHSA’s definition of SMI and was not used in the estimation of SMI.

Table 1. Diagnostic Modules in the 2008-2012 MHSS SCID

Mood Disorders Past Year Major Depressive Episode* Lifetime Major Depressive Episode Past Year Manic Episode* Lifetime Manic Episode Dysthymic Disorder*	**Past Year Eating Disorders** Anorexia Nervosa* Bulimia Nervosa* **Past Year Impulse Control Disorders** Intermittent Explosive Disorder*
Past Year Psychotic Disorders Psychotic Screen* **Past Year Anxiety Disorders** Posttraumatic Stress Disorder* Panic Disorder with and without Agoraphobia* Agoraphobia without History of Panic Disorder* Social Phobia* Specific Phobia* Obsessive Compulsive Disorder* Generalized Anxiety Disorder*	**Past Year Substance Use Disorders** Alcohol Abuse Alcohol Dependence Non-Alcohol Substance Abuse Non-Alcohol Substance Dependence **Past Year Adjustment Disorders** Adjustment Disorder*

MHSS = Mental Health Surveillance Study; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition.

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

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

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

1. An open-ended overview module is part of the standard SCID designed to elicit information about the respondent’s diagnostic and treatment history and current status in a way that establishes some level of rapport between the clinical interviewer and the respondent.
2. The SCID screener instrument is a set of questions at the end of the overview section and is administered to all respondents. Its questions are taken from the body of the SCID and are the initial questions asked by the SCID for the disorders being assessed. These screening questions may help to reduce the potential effects of a “negative response bias” that may be especially problematic in the later sections of the SCID. Due to the structure of the SCID, respondents tend to notice that a “yes” answer to the initial probe question in a section results in follow-up questions, whereas a “no” answer results in a skip to the next section. Thus, some subjects will answer “no” to speed up the interview. By asking these questions up front and using the answers to these questions in the determination of whether a section should be skipped, response bias may be minimized.
3. With the DSM-IV Axis V GAF scale, the clinical interviewer was instructed to rate the respondent’s period of worst psychological, social, and occupational functioning during the past year.

4. A section is included for documenting the clinical interviewer's impressions of the interview situation, including (1) ratings of the respondent's level of privacy, cooperation, and comprehension; and (2) the overall validity of the interview data (any interview deemed by the clinical interviewer or clinical supervision team to be of questionable validity was discarded).

For more detail about these modules, see CBHSQ (2014a) and Colpe et al. (2010).

2.5 Estimation Methods in the MHSS

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

The primary objective of the MHSS analysis aimed to produce annual national estimates of SMI prevalence that are accurate for all adults and for adult subpopulations. Secondary objectives included producing estimates of other categories of mental illness defined by level of impairment, such as mild (or low) mental illness (LMI), moderate mental illness (MMI), and AMI. Table 2 defines these categories of mental illness, which are based on SCID disorder diagnoses and GAF scores. A respondent was coded positive for SMI if he or she was determined to have any of the mental disorders (not including developmental or substance use disorders) assessed in the MHSS SCID *and* had a GAF score of 50 or below. AMI, which is defined as having a mental disorder regardless of the level of impairment due to that disorder, is the category obtained by collapsing SMI, MMI, and LMI in Table 2 into a single category.

Table 2. Mental Illness Categories Defined by SCID Disorder Diagnosis and GAF Score: 2008-2012 MHSS Clinical Follow-Up Study

Mental Illness Category	SCID Disorder Diagnosis	GAF Score
AMI	1 or More	Not applicable
SMI	1 or More	GAF ≤ 50 (severe or worse impairment)
MMI	1 or More	50 < GAF ≤ 59 (moderate impairment) *
LMI	1 or More	59 < GAF (at most mild impairment)
No Mental Illness	None	Not applicable

AMI = any mental illness; GAF = DSM-IV Axis V Global Assessment of Functional Scale; LMI = mild (or low) mental illness; MHSS = Mental Health Surveillance Study; MMI = moderate mental illness; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

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

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

A logistic regression model was developed based on the 2008 clinical data to produce a predicted probability of having SMI for each clinical interview respondent. The independent variables (predictors) in the model were from the main NSDUH interview. The final "best" model included two predictors: Alternative Past Year K6 score and Alternative WHODAS score. A cut point was established among the fitted probabilities of having SMI based on the 2008 MHSS clinical data. Adults with probabilities at or above the cut point were predicted to have SMI and the rest were not. The cut point chosen was the value at which the weighted sum of false positives (respondents for which the model predicted SMI, but the clinical interview indicated no SMI) was close to the weighted sum of false negatives (respondents for which the model predicted no SMI but the clinical interview indicated they had SMI). The model parameters estimated from the clinical sample were applied to the full 2008 NSDUH adult sample (all predictors were in the main NSDUH interview) to produce a predicted probability of SMI and SMI indicator (using the cut point) for each adult respondent in 2008. The same model was applied to the full 2009-2011 NSDUH data as they became available for the purposes of producing estimates of SMI annually. Although this model was optimized to predict SMI, the SMI predicted probabilities were also used to predict MMI and LMI using different cut points (for more details about this model, see CBHSQ, 2014a).

With the accumulated MHSS clinical data collected from 2008-2012, SAMHSA determined that the 2008 model had some important shortcomings that had not been detected in the original model fitting due to the small number of respondents in the 2008 clinical sample. Specifically, the 2008 model substantially overestimated SMI and AMI among young adults (i.e., ages 18 to 25) in NSDUH relative to young adults in the clinical interview data. In addition, improvements were needed in the weighting procedures for the MHSS clinical data to better account for undercoverage and nonresponse (i.e., because only NSDUH respondents who answered their surveys in English were eligible for the clinical follow-up and because persons with mental illness appeared to be more likely to participate in the follow-up). Furthermore, SAMHSA wanted to determine if the model could be improved by adding more mental health-related predictor variables. Therefore, using the combined 2008-2012 clinical data, SAMHSA fit a more accurate model for the 2012 estimates with revised weights (subsequently referred to as the "2012 model" and "2012 estimation methods") and additional predictors (see Section 2.6 for more details). In particular, to reduce bias at the subpopulation level and improve prediction, additional mental health-related variables and an age variable were added in the 2012 model. In addition, to protect against potential coverage and nonresponse error, alternatives for the weights were applied to the clinical sample data for the model development. To provide consistent data for trend assessment, mental illness estimates for 2008-2011 using the new 2012 model were revised. See CBHSQ (2014a) for further details.

2.6 Evaluating Alternative Models

By the end of 2012, a nationally representative sample of 4,912 completed clinical assessments had been collected among respondents assigned to the WHODAS questions. This larger dataset was used to identify and evaluate alternative models of SMI to produce cut point estimators for SMI and AMI.

A variety of variables was considered as predictors for an alternative model. The following criteria were used to decide on the number and type of variables that could be included in the model. First, a limited number of covariates could reasonably be added to a logistic model based on clinical sample data having, at most, 100 effective degrees of freedom. The covariates would be added after collapsing strata in the NSDUH full adult sample (see Section 3.7 of CBHSQ (2014a) for more information) to ensure that no primary sampling unit (PSU) is empty (100 variance strata with two variance PSUs each). A maximum of 10 (i.e., $\approx \sqrt{100}$) variables were considered for inclusion in each model.[8]

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

Ultimately, the goals was to construct a parsimonious model that could be used annually to assess changes in SMI (and other categories of mental illness) within and between sociodemographic groups. A covariate related to a domain of interest would tend to fix the relationship between SMI and that domain.[9] For example, if being employed resulted in an estimated 2% decrease in the odds of having SMI when all other things were equal, then treating that 2% decrease as fixed over time would impede the measurement of any changes in the relationship between SMI and employment. Therefore, in most instances variables closely related to domain-membership indicators were not considered for inclusion into the model as covariates.

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

Different combinations of the K6 and WHODAS items and scores were evaluated as predictors in a variety of models for SMI, but none led to meaningful reductions in the error rate when compared to the alternative K6 and WHODAS scores used in the 2008 model. In comparison, the addition of two variables from the main NSDUH interview, serious thoughts of suicide and the experience of a MDE during the past year, noticeably decreased the error rate and were included in a (potential) new model of SMI.

The second metric used to evaluate the models was a measure of bias. A cut point estimator for SMI prevalence is based on a model, which is as good as the assumptions on which it is based. As a result, unlike a model-free direct estimator computed directly from the clinical sample, a cut point estimator can be systematically biased. The bias measure for a SMI prevalence estimate can be measured among all adults or for domains. The bias measure for a domain was defined as the difference between the weighted proportions across the clinical sample within the domain of respondents predicted to have SMI and those actually diagnosed with SMI (equal to the difference

---

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

[9] The domains of interest were sex, age group, race/Hispanic origin, region, county type, employment level, education, whether the adult received mental health treatment, whether the adult had health insurance, and the adult's household income in relation to the poverty threshold. Definitions of these domains can be found in CBHSQ (2014a).

between the false positive and false negative rates in the domain). Under the null hypothesis that no bias exists in the domain, this bias measure would not be significantly different from zero.

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

Bias in estimates of SMI was investigated for a number of domains of interest (i.e., sex, age group, race/Hispanic origin group, region, county type, employment level, education, whether the adult received mental health treatment, whether the adult had health insurance, and the adult's household income in relation to the poverty threshold). To do this, a model of SMI was fit on the entire WHODAS-assigned MHSS clinical sample from 2008-2012. Predictor variables in this model of SMI included past year K6, WHODAS, MDE, and suicidal thoughts. Results of the investigation indicated that SMI estimates within certain age groups were biased (CBHSQ, 2014a).

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

The age variable that was ultimately added to the SMI model (termed AGE1830) was recoded from a continuous age variable for adults and is coded as either 12 or the difference between the respondent's age and 18, whichever was smaller. The variable increased as the respondent aged from 18 to 30, but then leveled off at 12 after age 30. Adding this age variable led to a cut point that both equalized false positives and false negatives for all adults and roughly within adult age groups.

After evaluating a wide variety of model specifications for both SMI and AMI that focused on finding a model with minimum values of bias measures and error rates, the final model included the following predictor variables: alternative past year K6 and WHODAS scores, past year MDE and suicidal thoughts, and AGE1830.

2.7 2012 NSDUH Prediction Model

The 2012 NSDUH prediction model provides national estimates of SMI and AMI based on data from the 2008A (i.e., Sample A of 2008 was assigned to the WHODAS questions) to 2012 NSDUHs for adults age 18 or older (CBHSQ, 2014a).[10] Specifically, π represents the probability that an adult has SMI. The 2012 NSDUH prediction model is expressed as—

[10] A separate model was developed for respondents assigned to the Sheehan Disability Scale (SDS) in 2008B (i.e., Sample B of 2008).

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

where $\hat{\pi}$ is the estimated probability that an adult had SMI, and the predictor variables are defined as—

X_k = *Alternative Past Year 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 of 0 to 8

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

X_m = *Past Year MDE on NSDUH*: coded as 1 if the criteria for past year MDE were met;[11] coded as 0 otherwise

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

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

$$\text{SMIPP\_U} = 1/\left(1+\exp\left[-\begin{pmatrix}-5.9726644 + 0.0873416 * X_k + 0.3385193 * X_w + \\ 1.9552664 * X_s + 1.1267330 * X_m + 0.1059137 * X_a\end{pmatrix}\right]\right).$$

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 past year AMI (AMIYR_U = 1). Otherwise, the respondent was predicted as not having past year AMI (AMIYR_U = 0).[12]

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

[12] Before the 2012 NSDUH prediction model had been developed, SMI predicted probabilities, SMI and AMI cut point estimates were produced from the 2008 NSDUH prediction model, and the corresponding variables were called SMIPP, SMIYR, and AMIYR. These variables are not comparable with their analogues produced from the 2012 SMI model (i.e., the variables SMIPP_U, SMIYR_U, and AMIYR_U).

2.8 Estimates of SMI and AMI for the 2008-2012 Combined Years

Table 3 displays direct estimates of SMI and AMI across several demographic domains (i.e., subpopulations) based on the SCID from the clinical data for the 2008-2012 combined years. The table also shows corresponding model-based cut point estimates of SMI and AMI based on the final 2012 model applied to the adult NSDUH data for the 2008-2012 combined years. Tests of significance were also conducted between the two sets of estimates, and statistically significant differences at the .05 and .01 levels are displayed where applicable. Only one difference was statistically significant: the AMI cut point estimate was statistically significant compared to the AMI direct estimate in the South ($p < .01$).

Table 3. NSDUH Estimates of SMI and AMI Based on 2012 NSDUH Prediction Model, by Demographic Domain: 2008A-2012

	SMI Direct Estimate (MHSS)		SMI Cut Point Estimate (NSDUH)		AMI Direct Estimate (MHSS)		AMI Cut Point Estimate (NSDUH)	
Demographic Variable	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
Total	3.9%	0.284	3.9%	0.071	18.0%	1.058	18.1%	0.149
Sex								
Male	3.0%	0.374	2.9%	0.088	14.5%	1.211	14.3%	0.198
Female	4.8%	0.425	4.8%	0.107	21.3%	1.662	21.5%	0.217
Age								
18-25	3.8%	0.546	3.9%	0.082	21.2%	2.084	18.6%	0.171
26-34	4.3%	0.625	5.0%	0.175	19.5%	1.935	21.8%	0.348
35-49	5.7%	0.658	5.0%	0.138	20.5%	1.628	20.4%	0.265
50 or Older	2.7%	0.428	2.8%	0.119	14.8%	2.033	15.0%	0.271
Race/Hispanic Origin								
Non-Hispanic White	4.4%	0.357	4.2%	0.088	18.7%	1.000	19.0%	0.181
Non-Hispanic Black	3.3%	0.646	3.1%	0.151	15.0%	2.421	16.7%	0.410
Hispanic	2.0%	0.696	3.1%	0.173	18.9%	5.080	15.3%	0.402
Non-Hispanic Other	4.1%	1.168	3.2%	0.278	15.1%	3.108	16.8%	0.591
Region								
Northeast	2.8%	0.473	3.6%	0.140	19.1%	3.717	17.5%	0.322
Midwest	4.2%	0.484	4.2%	0.125	17.3%	1.442	18.4%	0.273
South	3.7%	0.486	3.7%	0.121	16.8%	1.350	17.8%[b]	0.249
West	5.0%	0.793	4.1%	0.161	19.7%	2.021	18.5%	0.355
County Type								
Large Metro	3.8%	0.414	3.6%	0.097	19.5%	1.779	17.4%	0.218
Small Metro	4.1%	0.506	4.1%	0.117	16.5%	1.323	18.8%	0.242
Nonmetro	4.0%	0.582	4.2%	0.186	16.2%	1.622	18.8%	0.362
Received Mental Health Services								
Yes	18.8%	1.681	18.5%	0.385	53.0%	2.822	53.9%	0.517
No	1.5%	0.172	1.5%	0.046	12.4%	1.120	12.3%	0.137
Employment								
Full Time	2.4%	0.305	2.8%	0.077	14.7%	1.043	15.2%	0.178
Part Time	4.3%	0.640	4.0%	0.167	20.3%	2.254	19.7%	0.388
Unemployed	5.6%	1.147	6.4%	0.321	21.2%	3.569	23.7%	0.601
Other	6.2%	0.752	5.2%	0.164	22.4%	2.806	21.1%	0.324
Education								
Less than High School	5.7%	1.071	4.0%	0.163	26.0%	5.262	19.6%	0.406
High School Graduate	4.0%	0.570	4.1%	0.136	17.9%	1.653	17.8%	0.267
Some College	4.1%	0.460	4.4%	0.138	15.6%	1.313	19.5%	0.293
College Graduate	2.9%	0.458	3.1%	0.119	16.9%	1.517	16.2%	0.273
Poverty Level[1]								
< 100% Threshold	9.0%	1.373	7.0%	0.237	25.3%	2.693	25.6%	0.425
100% to 199% Threshold	5.6%	0.840	4.8%	0.163	24.5%	3.868	21.1%	0.350
≥ 200% Threshold	2.6%	0.253	3.0%	0.078	15.0%	0.952	15.7%	0.174
Health Insurance								
Yes	3.6%	0.298	3.6%	0.077	17.1%	1.168	17.4%	0.160
No	5.7%	0.848	5.1%	0.175	22.8%	2.561	21.2%	0.352
Sample Size	4,912		206,500		4,912		206,500	

AMI = any mental illness; CBHSQ = Center for Behavioral Health Statistics and Quality; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SE = standard error; SMI = serious mental illness.

[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. Excludes adults ages 18 to 22 in a college dormitory. The sample sizes of this variable for the MHSS was 4,834 and 202,200 for NSDUH.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100.

Datasets = 2008A-2012 adult NSDUH data for model-based cut point 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: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

3. Summary of SPI

3.1 Purpose and History of SPI

BJS has periodically fielded SPI, a national, omnibus, cross-sectional survey of state and federal prisoners incarcerated in the United States. The state prisoner survey was first fielded in 1974, and the federal prisoner component was added in 1991.[13] The 2016 iteration of the SPI was the most recent administration of the study, representing the seventh national study of its kind among state prisoners and the fourth among federal prisoners.[14] The SPI's primary objective is to produce national statistics of the state and federal prison populations across a variety of domains, including but not limited to, individual characteristics, current offense and sentence, incident characteristics, criminal history, socioeconomic characteristics, family background, drug and alcohol use and treatment, mental and physical health and treatment, and facility programs and rule violations.[15] The data are critical to producing national statistics that provide an understanding of the composition of the U.S. prison population and the changes that occur over time; factors related to the changes observed, including the impact of corrections policy and practice reforms; the risk prisoners pose to correctional agencies and for recidivism; and the challenges prisoners face when they return to the community.

Beginning in 1997, the survey was administrated via computer-assisted personal interviewing (CAPI), while prior to this year, paper-and-pencil interviewing (PAPI) was used. Through a competitive bidding process, BJS selected RTI International to conduct the data collection for the 2016 SPI, which yielded about 25,000 completed interviews. Previous iterations of the SPI were conducted by the U.S. Census Bureau.

3.2 Survey Design and Response Rates of 2016 SPI

BJS's 2012 Census of State and Federal Adult Correctional Facilities (CCF, formerly CSFACF) served as the basis for the 2016 SPI universe. For each facility listed, the 2012 census enumerated the government type (state or federal), facility type (confinement-based or community), facility location, facility operator (public or private), number of prisoners held by sex, and whether the facility's main function was to provide mental health services. Between the completion of the CCF and prior to July 2014, when the SPI sample of prisons was selected, the 2012 CCF was updated to account for known changes in facilities. The types of changes included (1) adjusting the population size of a facility to account for a planned change in population, (2) removing facilities that planned to close by the time the survey was fielded, and (3) adding new facilities that were operational prior to selecting the SPI sample.

---

[13] Since the federal prisoner component was added in 1991, the survey has been conducted among state and federal prisoners concurrently through a single administration using the same questionnaire and data collection protocols.
[14] With the 2016 administration, BJS changed the name of the survey from the Survey of Inmates in State and Federal Correctional Facilities to SPI.
[15] The 2016 SPI was also designed to produce subnational estimates of jurisdictions with large prison populations, or those holding 100,000 or more prisoners. Some of the previous SPI iterations were also designed to produce subnational estimates for a small number of large prison population jurisdictions.

The 2016 SPI sample was a stratified two-stage design in which state and federal correctional facilities were selected in the first stage and prisoners within sampled facilities were selected in the second stage. The SPI sampling population consisted of two frames: facilities that housed (1) male and (2) female prisoners. Facilities that housed a combination of male and female prisoners were listed on both frames. Within each frame, facilities were stratified based on government type (state or federal) and, among state prisons, geographic area. The geographic area stratum split out states (three) with 100,000 or more prisoners housed.

An initial sample of 465 facilities was selected, made up of a 415-facility main sample and a 50-facility reserve sample. The main sample assumed an 85% participation rate (unweighted), and included: (1) an initial, random replicate of 355 facilities (main sample); (2) an additional, random replicate of 60 facilities of which a random subsample of 30 facilities was released into the field after data collection began; and (3) a 10-facility set aside as an oversample of federal prisons. The 50-facility reserve sample consisted of state prisons to be released if the first-stage participation rate dropped below 85%.

The sample of prisons was allocated across strata in a multi-step process. First, a size measure was developed for each facility. Given the frame, the base size measure was the number of male or female prisoners housed in the facility. The size measure was increased by a factor of 3.5 for female state facilities and 2.4 for female federal facilities. The size measure was further increased by a factor of three for facilities whose primary function was to provide mental health services. Second, the sample was allocated across the two frames proportionally based on the sum of the size measures. Third, within each frame, the sample of prisons was allocated across government type (state or federal) and geographic area (for state prisons only) proportionally based on the total size measure within each stratum.

Within each stratum, facilities were randomly selected with probability proportionate to their size measure. In the first step of the sampling process, self-representing (SR) facilities (i.e., facilities with an expected selection rate greater than 1) were automatically selected first, and the within-stratum sample size was reduced accordingly. Next, the combined main and reserve samples were selected after implicitly stratifying facilities by government type, whether the facility's primary function was mental health services, U.S. Census region, and state (state prisons only). In addition, a 10-facility oversample was proportionately allocated across the federal stratum during this step. Third, the reserve sample was selected via a systematic sampling after facilities were sorted by sex housed, population size, government type, geographic area (state prisons only), and facility operator. Finally, among the 415 main sample facilities, a random subset of 385 facilities (initial replicate of 355 facilities plus the subsample replicate of 30 facilities) was selected in a similar manner to the reserve sample and released into the field. It was determined during the first half of data collection, fewer than 415 facilities were necessary to obtain the target number of interviews, given the high first-stage response rate.

SPI was designed to be self-weighting within each stratum. As such, within each sampled facility, a fixed respondent sample size of 64 respondents in a state facility and 80 respondents within a federal facility was targeted. Within each facility, a response rate of 70% was assumed, yielding a starting sample size of 92 prisoners in state facilities and 115 prisoners in federal facilities. The exception to this sample size occurred when the targeted sample size exceeded 75% of the facility's total population. When this occurred, the sample size was set at 75% of the facility population.

In the second stage of selection, prisoners were sampled differently depending on the jurisdiction of the facility. Within state facilities, prisoners were selected via a simple random sample. Within federal facilities, prisoners were first stratified by their offense type (i.e., drug offense or non-drug offense because almost half of all federal prisoners were serving a sentence for a drug offense; Carson, 2014). Then, within each of the two strata, a simple random sample of prisoners was selected. Federal prisoners in the non-drug offense stratum were oversampled by a factor of 1.5 (e.g., if 30% of a facility's population consisted of prisoners with non-drug offenses, then 45% of the sample would be comprised of prisoners with non-drug offenses).

A total of 24,848 prisoners—20,064 state prisoners and 4,784 federal prisoners—in 364 state and federal facilities (306 state and 58 federal prisons) participated in the 2016 SPI. The first stage response rate (i.e., the response rate among selected prisons) was 98% (98% among state prisons and 100% among federal prisons). The second stage response rate (i.e., the response rate among prisoners within selected facilities) was 70% (69% among state prisoners and 73% among federal prisoners).

3.3 Mental Health Items in the 2016 SPI Questionnaire

The 2016 SPI questionnaire contains several mental health items that could be recoded into potential predictor variables consistent with NSDUH variables for SMI models:

- **MH1 – MH6**: correspond exactly with the six K6 items on a past month time scale
- **MH7**: asks "Have you **ever** been told by a medical doctor or a mental health professional, such as a psychiatrist or psychologist that you had"—
  - **MH7b**: "a depressive disorder?"
  - **MH7e**: "an anxiety disorder, such as panic disorder or obsessive-compulsive disorder, also known as OCD?"

The items above could be recoded into variables representing past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety.

3.4 Mental Health Variables Common to 2008-2012 NSDUH and 2016 SPI

The items in the 2016 SPI that can be recoded into variables representing past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety have similar counterparts in the 2008-2012 NSDUH.

The 2008-2012 NSDUH variables representing these measures include the following:

- **K6SCMON**: past month K6 score
- **K6SCMON2**: alternative past month K6 score (K6 score of less than 8 recoded as 0; past year K6 score of 8 to 24 recoded as 1 to 17)

- **DEPRSLIF_U**: reported having had depression in lifetime (coded as 1 if "yes"; coded as 0 otherwise [i.e., includes "no" and missing values])
- **ANXDLIF_U**: reported having had anxiety in lifetime (coded as 1 if "yes"; coded as 0 otherwise [i.e., includes "no" and missing values]).

Missing values among the K6SCMON and K6SCMON2 scores in the NSDUH were imputed as zeroes. An evaluation of item nonresponse (i.e., missing) rates among mental health variables in the 2010 NSDUH indicated that the K6SCMON score (and equivalently, K6SCMON2 score) had a weighted nonresponse rate of 1.0%; the K6SCMON score was considered missing if one or more of the six K6 items constituting the score was missing.[16]

The 2016 SPI variables representing the same measures were created as follows:

- **K6SCMON**: past month K6 score
- **K6SCMON2**: alternative past month K6 score (K6 score of less than 8 recoded as 0; past year K6 score of 8 to 24 recoded as 1 to 17)
- **DEPLIF**: reported having ever had a depressive episode [MH7b] (coded as 1 if "yes"; coded as 0 if "no")
- **ANXLIF**: reported having ever had an anxiety disorder [MH7e] (coded as 1 if "yes"; coded as 0 if "no").

Missing values among the 2016 SPI variables listed above were not imputed. The unweighted rates of missing values among the SPI sample were small: 1.2% for K6SCMON and K6SCMON2; 0.9% for DEPLIF; and 0.8% for ANXLIF. A variable was considered missing if one or more of the items constituting the variable was missing.

[16] See Chapter 9 in SAMHSA's Evaluation of Imputation Methods for the National Survey on Drug Use and Health.

4. Development of Method to Estimate SMI Among Prisoners

4.1 General Estimation Approach, Assumptions, and Criteria for Choosing and Ranking Models

Ideally, the estimation of SMI among prisoners would have followed a process similar to that of the MHSS in the NSDUH. In other words, a subsample of prisoners would have been selected from the 2016 SPI for participation in a clinical psychiatric interview to determine a gold-standard assessment of SMI. Then, for this subsample, alternative models and cut points could have been developed by matching the gold-standard measure of SMI with predictor variables derived from mental health questions in the SPI. However, in the absence of this ideal situation, a more circuitous analytical approach needed to be followed by drawing on information from the 2008-2012 NSDUHs and the 2016 SPI in six steps as follows.

4.1.1 Identification of Criminal Justice Subsample in NSDUH

Because no subsample of prisoner respondents with gold-standard assessments of SMI exists within the 2016 SPI, the respondents in the 2008-2012 MHSS clinical sample were used. The initial models and cut points potentially to be used for the prison populations were developed within the MHSS. However, for this approach to produce reasonably accurate estimates of mental illness for the prison populations, instead of using the entire household sample, a subsample of NSDUH respondents with characteristics most similar to those of prisoners was identified as the basis of modeling. This subsample consisted of adult respondents who had reported themselves to be *parolees*, *probationers*, or *arrestees* (i.e., respondents who had been arrested on one or more occasions) during the past 12 months. This nationally representative subsample is referred to as the *criminal justice* (CJ) subsample. In the main adult NSDUH survey for 2008-2012, a total of 15,400[17] respondents belonged to the CJ subsample (6.7%), and in the MHSS for 2008-2012, the size of the CJ subsample was 316 (5.6%). The full 2008-2012 MHSS sample (*n* = 5,653) is a representative sample of the full 2008-2012 NSDUH adult sample (*n* = 229,600), and the 316 CJ cases in the MHSS subsample are expected to constitute a representative sample of the 15,400 CJ cases from the main NSDUH because, from 2008 to 2012, the MHSS clinical sample was randomly selected from the main adult NSDUH. Table 4 compares demographic and CJ status characteristics among the CJ subsamples of the MHSS and NSDUH. The characteristics examined generally compare reasonably well between the two CJ subsamples. In particular, among the MHSS subsample, 54.9% were probationers, 18.1% were parolees, and 66.0% were arrestees; and among the main NSDUH adult sample, 55.0% were probationers, 17.3% were parolees, and 64.3% were arrestees. Because overlap exists among probationers, parolees, and arrestees, the sum of their percentages exceeds 100%. The estimated size of the CJ population based on the MHSS is 8,699,000 and 9,248,000 for NSDUH. Studies assessing the NSDUH estimates of probationers, parolees, and arrestees have generally found these estimates consistent with those published by BJS,

[17] Due to standard CBHSQ disclosure limitation protocols, certain unweighted NSDUH sample sizes were rounded to the nearest 100.

after accounting for the differing definitions and population coverage across data sources (Feucht & Gfroerer, 2011; Glasheen et al., 2012; Lattimore et al., 2014).

Table 4. Demographic, Mental Health, Substance Use, and Criminal Justice Characteristics Among the Criminal Justice in the MHSS and NSDUH, 2008-2012

Demographic/Criminal Justice Variable	MHSS Estimate	MHSS SE	NSDUH Estimate	NSDUH SE
Total	100.0%		100.0%	
Sex				
Male	77.1%	4.68	72.0%	0.62
Female	22.9%	4.68	28.0%	0.62
Age				
18-24	26.9%	4.57	30.6%	0.54
25-34	25.4%	4.61	29.3%	0.62
35 or Older	47.7%	6.87	40.1%	0.75
Race/Hispanic Origin				
Non-Hispanic White	59.1%	6.24	55.7%	0.75
Non-Hispanic Black	15.6%	4.16	21.0%	0.66
Hispanic	16.4%	6.25	19.0%	0.63
Non-Hispanic Other	8.9%	3.24	4.2%	0.25
Education				
Less than High School	34.2%	6.48	31.1%	0.70
High School Graduate	32.9%	4.74	36.6%	0.68
Some College	24.5%	3.93	24.3%	0.60
College Graduate	8.4%	3.12	8.1%	0.45
Marital Status				
Married	31.1%	6.60	22.2%	0.65
Widowed	3.9%	3.05	2.0%	0.26
Divorced or Separated	12.4%	3.23	19.4%	0.67
Never Married	52.5%	6.17	56.4%	0.73
Military Service				
Yes	4.8%	1.79	8.4%	0.53
No	95.2%	1.79	91.6%	0.53
Received Mental Health Treatment in Past Year				
Yes	17.8%	4.11	21.6%	0.62
No	82.2%	4.11	78.4%	0.62
Illicit Drug or Alcohol Abuse or Dependence in Past Year				
Yes	41.6%	6.09	37.7%	0.68
No	58.4%	6.09	62.3%	0.68
Criminal Justice Status*				
Probationer	54.9%	5.65	55.0%	0.74
Parolee	18.1%	4.42	17.3%	0.56
Arrestee	66.0%	5.20	64.3%	0.72
Population Estimate (in thousands)	8,699	1,002	9,248	148
Sample Size	316		15,400	

CBHSQ = Center for Behavioral Health Statistics and Quality; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SE = standard error.

*Details do not sum to totals due to overlap among probationers, parolees, and arrestees.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100.

Datasets = 2008-2012 adult NSDUH data for NSDUH estimates, and 2008-2012 MHSS clinical data for MHSS estimates.
Analysis weight = ANALWT/5 for NSDUH estimates, and MHFNLWGT for MHSS estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

4.1.2 Fitting SMI Models in CJ Subsample of MHSS

Different iterations of weighted logistic regression models were fitted in the CJ subsample of the MHSS (n = 316). For these models, the response variable was the binary gold-standard SMI variable, and predictor variables included (some version of) past month K6 score, lifetime anxiety, lifetime depression (see Section 3.4), and age variables similar to the variable AGE1830 used in the 2012 NSDUH model (see Section 2.7). These variables were selected because similar versions can be created from items contained in the 2016 SPI questionnaire, a necessary condition for applying the model parameters to the SPI data to generate predicted probabilities for SPI respondents. Initial crosstabulations demonstrated strong correlations between each of the three mental health variables and predicted SMI in the full CJ sample (Table 5) and gold-standard SMI within the MHSS CJ subsample (Table 6), suggesting that they would all be good predictors in multiple regression models. Tables 5 and 6 are presented and discussed in further detail in Section 4.2.

Because these fitted regression models represented the first step in a multi-step estimation process, the model fit statistics (e.g., Wald or t statistics of beta estimates and associated p-values) were not useful in identifying the best models with which to proceed. As noted in Section 2.6, in the MHSS, the SMI models based on the different predictor variables were evaluated using mainly two metrics: the overall error rate and domain-level bias of the cut point estimators. A large set of candidate models was selected to proceed to the next step where the overall error rate and domain-level bias of the cut point estimators were determined.

Table 5. SMI Estimates, by Mental Health Indicator and Criminal Justice Status Among Adults 18 or Older, 2008-2012 NSDUH

	Criminal Justice Status														
	All Adults			On Probation or Parole†			Arrestee, Not on Probation or Parole			Arrestee, on Probation or Parole‡			Not Arrestee, Not on Probation or Parole		
Mental Health Indicator	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**
Total	3.9	%	0.07	10.0	%	0.59	9.5	%	0.67	9.8	%	0.45	3.6	**	0.07
Past Month K6 Score															
< 8	0.8	%	0.03	2.1	%	0.40	1.4	%	0.26	1.8	%	0.27	0.7	**	0.03
8-12	11.4	%	0.31	13.9	%	1.45	13.2	%	1.77	13.6	%	1.12	11.2	**	0.32
> 12	41.7	%	0.71	42.4	%	2.39	39.6	%	3.13	41.2	%	1.90	41.7	%	0.77
Lifetime Depression															
No	1.5	%	0.04	4.6	%	0.46	5.3	%	0.56	4.9	%	0.36	1.4	**	0.04
Yes	20.3	%	0.38	34.4	%	2.28	32.1	%	2.66	33.5	%	1.76	19.5	**	0.39
Lifetime Anxiety															
No	2.3	%	0.05	5.7	%	0.48	6.0	%	0.62	5.8	%	0.38	2.2	**	0.05
Yes	21.1	%	0.46	36.6	%	2.51	33.6	%	2.74	35.5	%	1.86	20.1	**	0.47
Past Month K6 Score, Lifetime Depression, and Lifetime Anxiety															
< 8, No Lifetime Depression or Lifetime Anxiety	0.3	%	0.02	0.9	%	0.21	0.8	%	0.23	0.8	%	0.16	0.3	**	0.02
> 12, with Lifetime Depression and Lifetime Anxiety	66.5	%	1.17	69.4	%	3.90	*64.8*	*%*	*5.87*	67.7	%	3.33	66.3	%	1.28
Sample Size	229,600			8,800			6,600			15,400			213,000		

CBHSQ = Center for Behavioral Health Statistics and Quality; K6 = Kessler 6-item distress scale; NSDUH = National Survey on Drug Use and Health; SE = standard error.

*Difference between this estimate and estimate from comparison group † is statistically significant at the 0.05 level.

**Difference between this estimate and estimate from comparison group ‡ is statistically significant at the 0.05 level.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules.

Dataset = 2008-2012 adult NSDUH data.
Analysis weight = ANALWT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 6. Gold-Standard SMI Direct Estimates, by Mental Health Indicator and Criminal Justice Status Among Adults 18 or Older, 2008-2012 MHSS

	Criminal Justice Status														
	All Adults			On Probation or Parole†			Arrestee, Not on Probation or Parole			Arrestee, on Probation or Parole‡			Not Arrestee, Not on Probation or Parole		
Mental Health Indicator	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**
Total	3.9	%	0.27	*16.0*	%	*4.62*	*10.7*	%	*4.88*	13.9	%	3.29	3.5	**	0.25
Past Month K6 Score															
< 8	1.3	%	0.21	*5.9*	%	*3.03*	*8.4*	%	*6.20*	*7.1*	%	*3.31*	1.1	%	0.20
8-12	13.0	%	1.45	*17.9*	%	*11.12*	*12.1*	%	*7.29*	*15.8*	%	*7.14*	12.0	%	1.37
> 12	32.4	%	2.78	*40.3*	%	*9.68*	*30.2*	%	*11.38*	*38.4*	%	*8.39*	31.3	%	2.72
Lifetime Depression															
No	1.7	%	0.22	7.4	%	2.86	*9.6*	%	*5.32*	8.3	%	2.74	1.3	**	0.21
Yes	19.3	%	1.49	*61.5*	%	*8.70*	*18.7*	*	*10.24*	*46.6*	%	*9.08*	17.9	**	1.41
Lifetime Anxiety															
No	2.5	%	0.25	9.0	%	3.06	*9.9*	%	*5.36*	9.4	%	2.82	2.1	**	0.24
Yes	19.9	%	2.17	*53.0*	%	*13.36*	*16.2*	*	*10.07*	*39.9*	%	*10.72*	18.4	%	2.04
Past Month K6 Score, Lifetime Depression, and Lifetime Anxiety															
< 8, No Lifetime Depression or Lifetime Anxiety	0.8	%	0.20	*3.6*	%	*2.58*	*9.3*	%	*6.86*	*6.3*	%	*3.44*	0.6	%	0.18
> 12, with Lifetime Depression and Lifetime Anxiety	55.6	%	4.41	*70.9*	%	*13.46*	*62.7*	%	*26.56*	*70.1*	%	*12.59*	52.8	%	4.29
Global Assessment of Functioning Score															
≤ 50	88.2	%	3.68	*68.1*	%	*12.26*	*83.6*	%	*12.21*	*72.3*	%	*9.21*	*91.0*	%	*4.18*
Sample Size	5,653			195			121			316			5,323		

K6 = Kessler 6-item distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SE = standard error.

*Difference between this estimate and estimate from comparison group † is statistically significant at the 0.05 level.

**Difference between this estimate and estimate from comparison group ‡ is statistically significant at the 0.05 level.

Note: Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules (see Appendix A for more details).

Dataset = 2008-2012 MHSS clinical data.
Analysis weight = MHFNLWGT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

4.1.3 Determining Cut Points for SMI Models in CJ Subsample of MHSS

The weighted logistic regression models were used to produce predicted probabilities of SMI that are theoretically continuous between 0 and 1. For each model, the next step was to dichotomize these predicted probabilities into a binary variable indicating predicted SMI status ("yes" or "no") by using a cut point (i.e., respondents whose predicted probabilities were at or above the cut point would be predicted to have SMI and all others would be predicted not to have SMI).

For a particular model, a cut point was determined to ensure that SMI estimates based on predicted SMI status were unbiased (i.e., equal to the prevalence estimate based on gold-standard SMI responses). There are two equivalent operational methods for achieving this cut point:

1. Select the cut point that results in the weighted false positive and false negative rates (defined in Section 4.1.4) being as equal as possible. The weights may prevent exact equality from being achieved.
2. Select the cut point that results in the SMI cut point estimate (i.e., weighted mean of predicted SMI status achieved after application of cut point) being as equal as possible to the weighted mean of the predicted probabilities. This is appropriate because the weighted mean of the predicted probabilities from a logistic regression model is equal to the weighted mean of the response variable (i.e., gold-standard SMI).

A cut point was determined for each of the SMI models selected using the MHSS sample. In each case, a cut point estimator was derived from the binary variable indicating predicted SMI status.

4.1.4 Examination of the SMI Cut Point Estimators

A receiver operating characteristic (ROC) analysis was conducted on the cut point estimators derived from the selected SMI models and associated cut points. In the analysis, the following ROC statistics were reported:

- *False positive* (FP) rate is the proportion of all respondents who were predicted to have SMI but did not have (gold-standard) SMI.
- *False negative* (FN) rate is the proportion of all respondents who were predicted not to have SMI but did have (gold-standard) SMI.
- *Total classification error* (TCE) rate is the sum of the FP and FN rates.
- *Sensitivity* is the proportion of respondents with SMI who were predicted to have SMI.
- *Specificity* is the proportion of respondents without SMI who were predicted not to have SMI.
- *Area under ROC curve* (AUC) based on predicted responses dichotomized by the cut point is the average of sensitivity and specificity.

In the MHSS, the SMI models based on the different predictor variables were evaluated using mainly two metrics: the TCE and domain-level bias (see Section 2.6). Similarly in this study, cut point estimators (and associated models) were then ranked by TCE rate (smaller is better), subject to having a small overall absolute bias (i.e., the absolute value of the difference between the FP and

FN rates). Although the cut points were determined by minimizing the absolute bias of the cut point estimators, it was not always possible to have a small resulting absolute bias. For example, models based on small sample sizes, variable weights, or with few predictor variables resulting in a small number of distinct predicted probabilities may lead to a scenario where a cut point defined at a particular predicted probability may result in too many false positives. However, if increased to the subsequent predicted probability, it would result in too many false negatives.

The TCE rate is almost a perfectly monotone decreasing function of the AUC (i.e., as the TCE rate goes down, the AUC will almost always increase). Therefore, the cut point estimator with the smallest TCE rate typically also has the largest AUC. An additional step in assessing cut point estimators was to examine if the bias was maintained at the demographic domain (i.e., subpopulation) level. A cut point estimator with a small absolute bias at the overall level does not necessarily result in small absolute bias at the domain level across demographic variables.

A subset of models was selected where the resulting cut point estimators ranked best according to the TCE rate. Due to the small sample size, no estimates of bias overall and at the sub-domain level were statistically different from zero. Nevertheless, cut point estimators with smaller estimates of absolute bias were preferred to those with larger estimates of absolute bias.

4.1.5 Extrapolation of SMI Models and Cut Points to Main NSDUH Survey

The CJ subsample in the MHSS is 316, and in the main adult NSDUH survey, the CJ sample size is 15,400. Therefore, if it is reasonable to extrapolate the selected SMI models and cut points to the CJ subsample of the main survey, then it is possible to use the much larger dataset to derive national estimates of SMI for the CJ population based on the resulting cut point estimators.

For a subsample of size *n* = 316, the model may extrapolate to the larger sample reasonably well (i.e., the cumulative distribution functions [CDFs] of the predicted probabilities are similar in the two samples). However, for small samples, the choice of the cut point may occur in a region of the CDF that is not smooth due to the small sample size of the subsample. In this instance, a minor shift in the cut point could cause a large change in the prevalence estimate, indicating instability. This would likely occur if the SMI cut point estimate differs considerably from the weighted mean of the predicted probabilities in the main NSDUH data. A *hybrid cut point* accounts for this error by applying the second operational method of determining the cut point (see Section 4.1.3) to the *main NSDUH data rather than to the MHSS subsample*. In other words, the hybrid cut point is selected such that the SMI cut point estimate is as equal as possible to the weighted mean of the predicted probabilities in the main NSDUH data. See CBHSQ (2015a) for technical details about the hybrid cut point estimator.

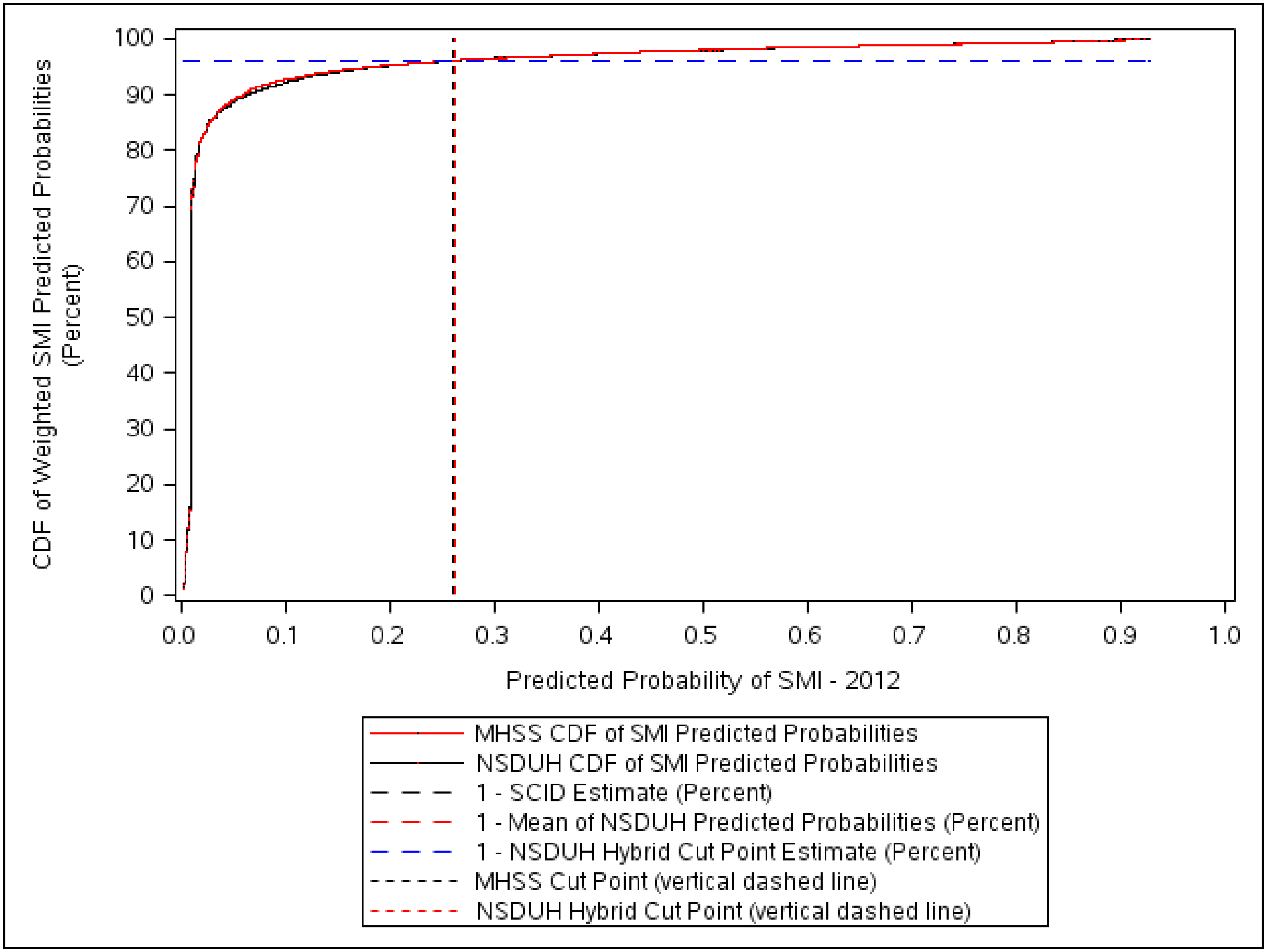
To illustrate the issue of extrapolating a model and cut point from a subsample of a larger sample to the larger sample itself, two figures are presented.

In Figure 1, the CDF of SMI predicted probabilities based on the 2012 NSDUH model developed in the 2008A-2012 MHSS (i.e., restricted to respondents assigned to the WHODAS questions; see Section 2.7 for more details) is presented in red. The CDF of SMI predicted probabilities based on the same model but applied to the full adult NSDUH data is presented in black. Due to the relatively large sample sizes of the 2008A-2012 MHSS (*n* = 4,912) and the 2008A-2012 adult NSDUH (*n* =

206,500), the CDFs are smooth and coincide closely. This indicates that the final SMI model developed in the MHSS *extrapolates* well to the adult NSDUH data. In the figure, three dashed lines indicate percentages of the following three functions of estimates: (1) black: 1 – SCID direct estimate; (2) red: 1 – mean of NSDUH predicted probabilities; and (3) blue: 1 – NSDUH hybrid cut point estimate. All three estimates almost perfectly coincide, so one dashed line appears visible. Finally, also in Figure 1 are the MHSS cut point plotted (black vertical dashed line) and the NSDUH hybrid cut point (red vertical dashed line). Again, because the two cut points are nearly identical, one vertical dashed line appears visible. These results indicate that not only does the final SMI model extrapolate to the larger NSDUH data well, but the cut point (and resulting cut point estimator) does as well. Note that the cut points meet the CDFs at the same point where the three functions of estimates do.

The results in Figure 2 paint a somewhat different picture. Figure 2 presents similar information to that presented in Figure 1, except that (1) predicted probabilities were based on the “final” SMI model selected in this study (discussed later in Section 4.6), and (2) only the CJ subset of both MHSS and NSDUH datasets was used, resulting in a substantial reduction in sample sizes. Recall that for the CJ subset, the sample sizes are 316 for the MHSS and 15,400 for the full NSDUH. Neither of the CDFs is smooth, and although they might coincide in some places, they clearly do not in others. Therefore, it is difficult to assess how well the model extrapolates to the adult NSDUH data. The sample size of the CJ subsample of the 2008-2012 NSDUH is 15,400, so the lack of smoothness of the NSDUH CDF (in black) appears to be due to the relative parsimony of the model rather than to sample size. A more parsimonious model, i.e., one with fewer variables, results in fewer distinct predicted probabilities, thus the CDF will be coarser than that of a model with more terms or levels within terms. The lack of smoothness of the MHSS CDF (in red) is likely also due to the relative parsimony of the same model, in addition to the small sample size (n = 316). Additionally, some difference exists between the MHSS and NSDUH hybrid cut points, suggesting that the MHSS cut point does not extrapolate well to the NSDUH data, due to two CDFs not overlapping in the neighborhood of the cut point. Therefore, the NSDUH hybrid cut point needs to shift to the right to allow the resulting cut point estimate to closely match the SCID direct estimate.

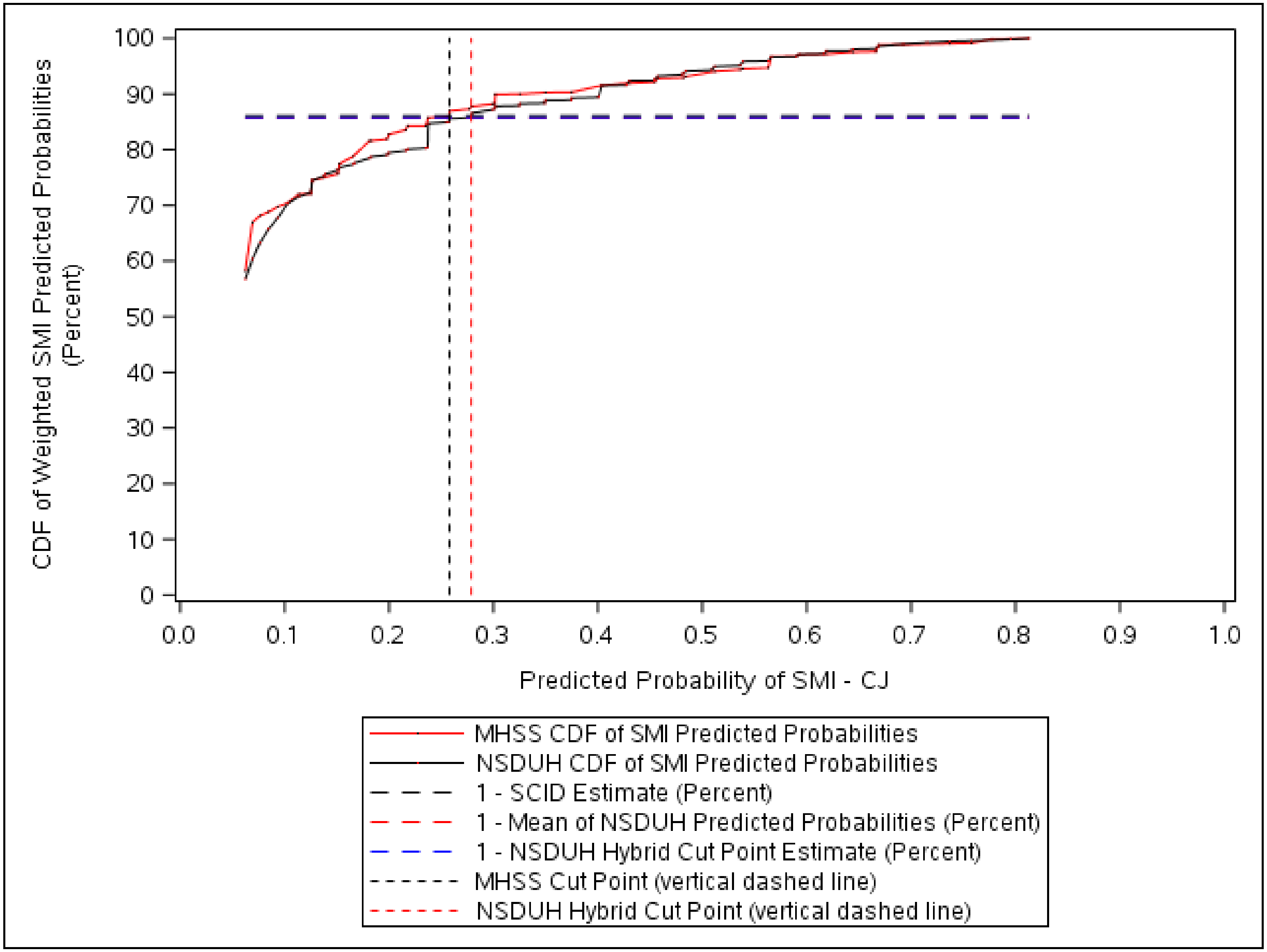
Figure 1. CDFs of Weighted SMI Predicted Probabilities Based on Final SMI Model, 2008A-2012



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

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

Figure 2. CDFs of Weighted SMI Predicted Probabilities Based on SMI Model for Criminal Justice Subpopulation, 2008-2012



CDF = cumulative distribution function; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

4.1.6 Final Extrapolation of SMI Models and Cut Points to 2016 SPI

Finally, the models and cut points selected in the previous steps were extrapolated to the prisoner data from the 2016 SPI to produce national estimates of SMI among state and federal prisoners. Because the extrapolation occurred across two different surveys (with different sampling and data collection methodologies) and populations (CJ within NSDUH versus prisoner within SPI), there was no direct method to test the success of the extrapolation. The resulting estimates of SMI may act as reasonable benchmarks if the assumptions hold regarding the transferability of the model, despite the differences in methodology and subpopulation. If feasible, future endeavors may want to consider collecting clinical information from a subsample of prisoners in SPI (e.g., similar to the MHSS).

4.2 Assessment of Data and Implications for Feasibility of Estimation Method

As noted in Section 4.1.1, the initial modeling analyses needed to be conducted within the MHSS because respondents in this subsample had both gold-standard assessments of SMI *and* responses to mental health items in the NSDUH questionnaire. However, this approach required the identification of the subsample of NSDUH respondents with similar characteristics as prisoners to serve as a proxy (i.e., the CJ subsample). Respondents who self-reported as parolees, or even probationers, appeared to be the best choice. However, the sample size of these respondents in the MHSS was 195, so respondents who self-reported as arrestees were added to this subsample to boost the sample size within the MHSS to 316.

Table 5 compares MHSS SMI cut point estimates applied to the following subsamples of the 2008-2012 adult NSDUH: (1) all adults; (2) those on probation or parole; (3) arrestees, but not on probation or parole; (4) arrestees, on probation or parole (i.e., the CJ subsample); and (5) all adults excluding the CJ subsample. The table also breaks out SMI estimates by several mental health indicators based on past month K6 score, lifetime depression, and lifetime anxiety.

Table 5 indicates that an estimated 9.8% of the CJ population has SMI based on the standard SMI prediction model used on the NSDUH for all adult respondents. This estimate is statistically significantly higher than the 3.6% estimated for the non-CJ population (i.e., the general population excluding the CJ population). Within the CJ population, 10.0% of parolees and probationers and 9.5% of arrestees (not on probation or parole) are estimated to have SMI—not a statistically significant difference. This pattern of results appears to hold across the mental health indicators shown in the table.

Table 6 compares gold-standard SMI direct estimates among the same subsamples of the 2008-2012 MHSS as Table 5 showed among the 2008-2012 NSDUH. Table 6 also breaks out SMI estimates by several mental health indicators based on past month K6 score, lifetime depression, lifetime anxiety, and GAF score.

Table 6 indicates that an estimated 13.9% of the CJ population has SMI, which is substantially higher than the 3.5% estimated for the non-CJ population. Within the CJ population, 16.0% of parolees and probationers and 10.7% of arrestees (not on probation or parole) have SMI. Although these estimates appear to be substantially different, both have large standard errors (4.6% of parolees and probationers and 4.9% of arrestees) that are not statistically significantly different and would have been suppressed according to the NSDUH precision-based suppression rules.[18] This pattern of results appears to hold across the mental health indicators shown in the table, although most of these estimates would be suppressed according to the NSDUH suppression rules.

The results in Tables 5 and 6 suggest the following course of action:

1. To have a sufficiently large sample size within the MHSS to continue SMI estimation for prisoners, all CJ parolees, probationers, and arrestees were included. In the smaller 2008-2012 MHSS, SMI estimates among arrestees are statistically significantly different from those among parolees or probationers only in the subdomains indicated by the presence of lifetime depression and lifetime anxiety. In the larger 2008-2012 NSDUH, their estimates are similar to those of parolees or probationers across all subdomains.
2. Parolees and probationers should be separated out for analyses in the final stages of this process to see if removal of the arrestees (but not probation or parole) affects the final model.

4.3 Preliminary Model Testing

Weighted logistic regression models were fit in the CJ subsample of the MHSS (n = 316). In these models, the response variable was the binary gold-standard SMI variable **SCID_SMI**, and predictor variables included various combinations of the four NSDUH variables described in Section 3.4 [i.e., K6SCMON (past month K6 score), K6SCMON2 (alternative past month K6 score), DEPRSLIF_U (lifetime depression), and ANXDLIF_U (lifetime anxiety)]. Also, prior experience from the MHSS indicated that AGE1830 was an important variable in the model to reduce the bias of estimates within age groups (CBHSQ, 2014a). Specifically, without this variable in the model, the SMI estimate in the 18-25 age group was substantially biased upward, and in the 35-49 age group it was substantially biased downward. After including the age variable, the bias in both of those age groups was substantially reduced. Therefore, variables representing age in various forms were also considered in this context. In particular, age variables of the following forms were considered:

- **AGE*n*** equals max(AGE – n, 0), where n has 2 digits. For example, AGE30 is equal to 0 for all ages up to age 30, then is equal to age minus 30 thereafter.
- **AGE*nm*** equals min(max(AGE – n, 0), m – n), where n and m each has 2 digits. For example, AGE1830 is equal to 0 for all ages up to age 18, then is equal to 30 minus 18 for ages between 18 and 30, and is equal to 12 thereafter.

Many models with various combinations of the predictor variables were run, but none included both K6SCMON and K6SCMON2 or more than one age variable in the same model. As noted in Section 4.1.2 (and Section 2.6), SMI models based on the different predictor variables were evaluated using mainly two metrics: the overall error rate and domain-level bias of the cut point estimators. Even so, the model fit statistics (e.g., Wald or t statistics of beta estimates and associated p-values) were still of interest, and models suggesting a poor fit were unlikely to be selected

[18] See Appendix A for details about the NSDUH suppression rules.

according to the two evaluation metrics. Therefore, a fairly large set of candidate models was selected to proceed to the next step.

For demonstration purposes, Table 7 displays the details of a couple of “representative” models. In both models, K6SCMON2 and DEPRSLIF_U were the only statistically significant predictor variables.

Table 7. Models K2A30 and K2DA: Predictor Variables, *DF*, Wald *F*, and Beta Statistics, 2008-2012 MHSS

Model K2A30						
Predictor Variable	***DF***	**Wald *F* Statistic**	**Wald *F* *p* Value**	**Beta**	**Beta SE**	**Beta *p* Value**
Intercept	1	46.5	0.000	-2.941	0.4314	0.000
K6SCMON2	1	5.9	0.018	0.112	0.0458	0.018
DEPRSLIF_U	1	8.3	0.006	1.497	0.5183	0.006
ANXDLIF_U	1	2.2	0.147	0.836	0.5679	0.147
AGE30	1	1.0	0.312	0.025	0.0243	0.312
Model K2DA						
Predictor Variable	***DF***	**Wald *F* Statistic**	**Wald *F* *p* Value**	**Beta**	**Beta SE**	**Beta *p* Value**
Intercept	1	36.7	0.000	-2.711	0.4474	0.000
K6SCMON2	1	5.6	0.022	0.109	0.0461	0.022
DEPRSLIF_U	1	9.8	0.003	1.544	0.4928	0.003
ANXDLIF_U	1	1.9	0.180	0.776	0.5703	0.180

CJ = criminal justice; *DF* = degrees of freedom; MHSS = Mental Health Surveillance Study; SE = standard error; SMI = serious mental illness.

Note: Sample size for both models is 316.

Response variable = SCID_SMI (gold-standard SMI).

Dataset = CJ subsample of 2008-2012 MHSS clinical data.
Analysis weight = MHFNLWGT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

4.4 ROC Analysis of SMI Models

A fairly large set of SMI models was selected to undergo the ROC analysis described in Section 4.1.4. For this analysis, a cut point was first chosen for each candidate model. The cut point was determined so that the predicted probabilities from the model could be dichotomized into a binary cut point SMI estimator that was (approximately) unbiased. Section 4.1.3 describes the methods for determining the cut point.

The models that underwent ROC analysis were ranked according to TCE rate (smaller is better), and simultaneously, models exhibiting smaller absolute bias were preferred. According to the ROC analysis, models that included the predictor variable K6SCMON2 (alternative past month K6 score) almost always performed better (in terms of TCE rate) than similar models that included K6SCMON (past month K6 score) (not shown). Therefore, all models that included the predictor variable K6SCMON were dropped from further consideration.

Table 8 displays the ROC analysis of a representative subset of models examined. The models in the table include the 10 top-ranked models (according to TCE rate) and models with different combinations of the mental health predictor variables to show the impact of excluding one or more mental health predictor variables.

The top 10 models in Table 8 all include the predictor variables K6SCMON2, DEPRSLIF_U, ANXDLIF_U, and some version of the AGE*n* or AGE*nm* variable. These models have TCE rates within a small range of 12.21 to 12.41 (the range of corresponding AUC values is 0.734 to 0.742). The absolute bias of these models is somewhat large and may have resulted from the small sample size. The only term that varies among these models is the age variable, suggesting that the SMI cut point estimator appears to be fairly robust to the actual age variable selected. Several models have identical ROC statistics, which indicates that such models result in identical cut point estimators.

The last four models in Table 8 exclude the age variable and up to two of the DEPRSLIF_U and ANXDLIF_U terms. The TCE rates for these four models are higher than the TCE rates for any of the models that include the age variable. For model K2DA, the TCE rate is 13.99. Models that exclude ages and one or both of the DEPRSLIF_U and ANXDLIF_U terms show even higher TCE rates (14.72 to 17.59).

The results in Table 8 suggest that all three mental health predictor variables (i.e., K6SCMON2, DEPRSLIF_U, and ANXDLIF_U) are important for the SMI cut point estimator, and some version of the age variable also appears somewhat important. The next step in this process was to gauge what effect the terms in the models have on the bias of the SMI cut point estimator at the *demographic domain level*.

Table 8. ROC Statistics of SMI Cut Point Estimates Based on Selected Candidate Models for Criminal Justice Population, 2008-2012 MHSS

Model	Sample Size	Pop Size (1,000s)	Cut Point	Gold-Standard Estimate	Cut Point Estimate	SE	False Positive Rate	False Negative Rate	Bias	TCE Rate	Sensitivity	Specificity	AUC
K2A30	316	8,699	0.245894	13.89	13.72	2.890	6.02	6.19	-0.1729	12.21	0.554	0.930	0.742
K2A3050	316	8,699	0.276321	13.89	13.30	2.890	5.84	6.43	-0.5913	12.28	0.537	0.932	0.734
K2A18	316	8,699	0.253753	13.89	13.58	2.891	5.98	6.30	-0.3137	12.28	0.547	0.931	0.739
K2A20	316	8,699	0.253077	13.89	13.58	2.891	5.98	6.30	-0.3137	12.28	0.547	0.931	0.739
K2A25	316	8,699	0.250791	13.89	13.63	2.891	6.02	6.28	-0.2586	12.30	0.548	0.930	0.739
K2A1845	316	8,699	0.264896	13.89	13.56	2.891	5.98	6.32	-0.3327	12.30	0.545	0.931	0.738
K2A2545	316	8,699	0.288997	13.89	13.56	2.891	5.98	6.32	-0.3327	12.30	0.545	0.931	0.738
K2A3040	316	8,699	0.284047	13.89	13.56	2.891	5.98	6.32	-0.3327	12.30	0.545	0.931	0.738
K2A3045	316	8,699	0.269146	13.89	13.56	2.891	5.98	6.32	-0.3327	12.30	0.545	0.931	0.738
K2A1850	316	8,699	0.275484	13.89	13.44	2.893	5.98	6.43	-0.4462	12.41	0.537	0.931	0.734
K2DA	316	8,699	0.257765	13.89	14.02	3.037	7.06	6.93	0.1286	13.99	0.501	0.918	0.710
K2D	316	8,699	0.324814	13.89	15.06	3.099	7.94	6.78	1.1649	14.72	0.512	0.908	0.710
K2A	316	8,699	0.276177	13.89	13.71	3.189	7.67	7.85	-0.1827	15.52	0.435	0.911	0.673
K2	316	8,699	0.212615	13.89	15.98	3.290	9.84	7.75	2.0846	17.59	0.442	0.886	0.664

AUC = area under curve (i.e., average of sensitivity and specificity); CJ = criminal justice; MHSS = Mental Health Surveillance Study; Pop = population; ROC = receiver operating characteristic; SE = standard error of cut point estimate; SMI = serious mental illness; TCE = total classification error.

Note: Bias = false positive rate – false negative rate; TCE rate = false positive rate + false negative rate. The estimated size of the CJ population based on the MHSS is 8,699,000 and 9,248,000 for NSDUH (see Section 4.1.1 for more details).

Response variable = SCID_SMI (gold-standard SMI).

Predictor variables included in the following models:

K2A30: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE30
K2A3050: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE3050
K2A18: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE18
K2A20: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE20
K2A25: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE25
K2A1845: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1845
K2A2545: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE2545
K2A3040: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE3040
K2A3045: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE3045
K2A1850: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1850
K2DA: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U
K2D: K6SCMON2 + DEPRSLIF_U
K2A: K6SCMON2 + ANXDLIF_U
K2: K6SCMON2

Dataset = CJ subsample of 2008-2012 MHSS clinical data.
Analysis weight = MHFNLWGT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

4.5 Assessing Subpopulation Bias for Reduced Set of Models

As noted in Section 4.1.4, a cut point estimator with a small absolute bias at the overall level does not necessarily mean that the absolute bias will be small at the domain (i.e., subpopulation) level across several demographic variables. Domain-level bias was a major consideration in the development of the 2012 NSDUH prediction model, which led to the addition of the AGE1830 variable in the model (CBHSQ, 2014a).

In this step, domain-level bias was assessed for the top 10 ranked models in Table 8 across the following demographic and geographic domains:

- sex (male, female)
- age (18 to 24, 25 to 34, 35 or older)
- race/Hispanic origin (non-Hispanic white, non-Hispanic black, Hispanic, non-Hispanic other)
- education (less than high school, high school graduate, some college, college graduate)
- marital status (married, widowed, divorced or separated, never married)
- military service ("yes," "no")
- received mental health services during the past year ("yes," "no")
- illicit drug or alcohol abuse or dependence during the past year ("yes," "no").

Table 9 displays the domain-level bias results for the 10 candidate models. Some models resulted in identical cut point estimators, so these models had the same domain-level bias. No models indicated any domain-level bias significant at the 0.05 level. However, this could be due to the small samples sizes of the domains (resulting in relatively large standard errors) rather than the point estimates of the bias itself. For example, for all models, the estimated bias for the "non-Hispanic other" race/Hispanic origin category was about 17%, and yet it was not significant at the 0.20 level. For all models, the point estimates of the bias for all sex, age group, and military service domains were relatively small (i.e., the absolute value of the bias was less than two percentage points). Of the 23 domain-level estimates among the models in Table 9, 16 (69.6%) would be suppressed due to low precision according to NSDUH suppression rules.

No particular model or models stood out as having lower domain-level bias results, so a subset was selected to extrapolate the models and cut points to the main NSDUH data. Section 4.6 discusses the results.

Four of the top 10 models in Tables 8 and 9 were selected for this step (i.e., models K2A30, K2A1845, K2A2545, and K2A1850) and two other models were also included (i.e., models K2DA and K2).[19] In previous MHSSs, the AGE1830 variable reduced the bias of estimates within age groups. Thus, four of the selected models include an age variable. By selecting multiple models that excluded an age variable, the impact of an age variable (positive or negative) on the TCE and bias

[19] See Table 8 for definitions of these models.

of cut point estimates could be compared. The latter two models without an age variable were selected.

Table 9. Bias of SMI Cut Point Estimates of 10 Candidate Models for Criminal Justice Population, 2008-2012 MHSS

Demographic Variable	n	SCID Estimate	Model: K2A30 Estimate	Model: K2A30 Bias		Model: K2A3050 Estimate	Model: K2A3050 Bias		Models: K2A18, K2A20 Estimate	Models: K2A18, K2A20 Bias		Model: K2A25 Esti-mate	Model: K2A25 Bias		Models: K2A1845, K2A2545, K2A3040, K2A3045 Esti-mate	Models: K2A1845, K2A2545, K2A3040, K2A3045 Bias		Model: K2A1850 Esti-mate	Model: K2A1850 Bias	
Total	316	13.9%	13.7%	-0.17		13.3%	-0.59		13.6%	-0.31		13.6%	-0.26		13.6%	-0.33		13.4%	-0.45	
Sex																				
Male	206	11.3%	11.1%	-0.14		10.9%	-0.36		11.1%	-0.21		11.1%	-0.14		11.1%	-0.21		10.9%	-0.35	
Female	110	*22.7%*	*22.4%*	*-0.30*		*21.3%*	*-1.37*		*22.0%*	*-0.67*		*22.0%*	*-0.67*		*21.9%*	*-0.76*		*21.9%*	*-0.76*	
Age																				
18-24	150	7.6%	9.0%	1.38		8.4%	0.78		8.5%	0.85		8.7%	1.06		8.4%	0.78		8.0%	0.36	
25-34	83	*16.3%*	*15.1%*	*-1.12*		*14.6%*	*-1.68*		*15.1%*	*-1.12*		*15.1%*	*-1.12*		*15.1%*	*-1.12*		*15.1%*	*-1.12*	
35 or Older	83	*16.2%*	*15.6%*	*-0.54*		*15.4%*	*-0.79*		*15.6%*	*-0.54*		*15.6%*	*-0.54*		*15.6%*	*-0.54*		*15.6%*	*-0.54*	
Race/Hispanic Origin																				
Non-Hispanic White	195	16.4%	16.3%	-0.14		15.9%	-0.55		16.1%	-0.35		16.1%	-0.29		16.1%	-0.35		16.1%	-0.35	
Non-Hispanic Black	40	*12.0%*	*20.2%*	*8.13*		*19.3%*	*7.23*		*20.2%*	*8.13*		*20.2%*	*8.13*		*20.2%*	*8.13*		*20.2%*	*8.13*	
Hispanic	50	*2.2%*	*3.0%*	*0.80*		*2.8%*	*0.56*		*2.9%*	*0.68*		*3.0%*	*0.80*		*2.8%*	*0.56*		*2.8%*	*0.56*	
Non-Hispanic Other	31	*21.8%*	*5.2%*	*-16.63*		*5.2%*	*-16.63*		*5.2%*	*-16.63*		*5.2%*	*-16.63*		*5.2%*	*-16.63*		*3.9%*	*-17.90*	
Education																				
Less than High School	78	*13.6%*	*6.3%*	*-7.36*	*	*5.7%*	*-7.93*	*	*6.2%*	*-7.46*	*	*6.3%*	*-7.36*	*	*6.1%*	*-7.52*	*	*6.1%*	*-7.52*	*
High School Graduate	117	13.7%	*21.0%*	*7.24*		*20.3%*	*6.56*		*20.6%*	*6.92*		*20.7%*	*6.98*		*20.6%*	*6.92*		*20.3%*	*6.58*	
Some College	96	*16.4%*	*17.6%*	*1.25*		*17.6%*	*1.25*		*17.6%*	*1.25*		*17.6%*	*1.25*		*17.6%*	*1.25*		*17.6%*	*1.25*	
College Graduate	25	*8.3%*	*4.2%*	*-4.13*		*4.2%*	*-4.13*		*4.2%*	*-4.13*		*4.2%*	*-4.13*		*4.2%*	*-4.13*		*4.2%*	*-4.13*	
Marital Status																				
Married	51	*17.1%*	*10.4%*	*-6.71*		*10.0%*	*-7.09*		*10.4%*	*-6.71*		*10.4%*	*-6.71*		*10.4%*	*-6.71*		*10.4%*	*-6.71*	
Widowed	6	*1.3%*	*1.9%*	*0.55*		*1.9%*	*0.55*		*1.9%*	*0.55*		*1.9%*	*0.55*		*1.9%*	*0.55*		*1.9%*	*0.55*	
Divorced or Separated	41	*22.2%*	*34.9%*	*12.68*		*34.8%*	*12.53*		*34.9%*	*12.68*		*34.9%*	*12.68*		*34.8%*	*12.53*		*34.8%*	*12.53*	
Never Married	218	11.0%	11.6%	0.60		11.0%	0.07		11.3%	0.33		11.4%	0.44		11.3%	0.33		11.1	0.12	

(continued)

Table 9. Bias of SMI Cut Point Estimates of 10 Candidate Models for Criminal Justice Population, 2008-2012 MHSS (continued)

Demographic Variable	n	SCID Estimate	Model: K2A30 Estimate	Model: K2A30 Bias		Model: K2A3050 Estimate	Model: K2A3050 Bias		Models: K2A18, K2A20 Estimate	Models: K2A18, K2A20 Bias		Model: K2A25 Estimate	Model: K2A25 Bias		Models: K2A1845, K2A2545, K2A3040, K2A3045 Estimate	Models: K2A1845, K2A2545, K2A3040, K2A3045 Bias		Model: K2A1850 Estimate	Model: K2A1850 Bias	
Military Service																				
Yes	15	*10.8%*	*12.5%*	*1.76*		*12.5%*	*1.76*		*12.5%*	*1.76*		*12.5%*	*1.76*		*12.5%*	*1.76*		*12.5%*	*1.76*	
No	301	14.0%	13.8%	-0.27		13.3%	-0.71		13.6%	-0.42		13.7%	-0.36		13.6%	-0.44		13.5%	-0.56	
Received Mental Health Treatment in Past Year																				
Yes	99	*42.8%*	*46.3%*	*3.46*		*45.9%*	*3.04*		*46.0%*	*3.15*		*46.3%*	*3.46*		*45.9%*	*3.04*		*45.9%*	*3.04*	
No	214	7.6%	6.7%	-0.96		6.3%	-1.38		6.6%	-1.06		6.6%	-1.06		6.6%	-1.06		6.4%	-1.20	
Illicit Drug or Alcohol Abuse or Dependence in Past Year																				
Yes	155	*15.7%*	*21.8%*	*6.10*		*21.4%*	*5.69*		*21.8%*	*6.02*		*21.8%*	*6.10*		*21.7%*	*5.97*		*21.7%*	*5.97*	
No	161	*12.6%*	*7.9%*	*-4.64*	*	*7.5%*	*-5.07*	*	*7.7%*	*-4.82*	*	*7.8%*	*-4.79*	*	*7.7%*	*-4.82*	*	*7.5%*	*-5.02*	*

CJ = criminal justice; MHSS = Mental Health Surveillance Study; n = sample size; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

*Bias is statistically significant from zero for $0.10 \leq p < 0.20$.

Note: Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. Bias = model-based cut point SMI estimate – SCID (i.e., gold-standard) SMI direct estimate.

Predictor variables included in the following models:

K2A30: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE30
K2A3050: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE3050
K2A18: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE18
K2A20: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE20
K2A25: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE25
K2A1845: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1845
K2A2545: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE2545
K2A3040: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE3040
K2A3045: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE3045
K2A1850: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1850

Dataset = CJ subsample of 2008-2012 MHSS clinical data.
Analysis weight = MHFNLWGT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

4.6 Assessing Models in Main 2008-2012 NSDUH

As noted in Section 4.1.5, the size of the CJ subsample in the MHSS is 316, but in the main adult NSDUH survey the CJ subsample size is 15,400. Therefore, if the selected SMI models and cut points can be extrapolated to the main survey, a much larger dataset can be used to derive national estimates of SMI based on the resulting cut point estimators. As discussed in Section 4.1.1, the full MHSS subsample is representative of the full NSDUH, so the 316 CJ cases in the MHSS subsample are a representative sample of the 15,400 CJ cases from the main NSDUH.

As noted in Section 4.5, four of the models in Table 8 were selected for this step (i.e., models K2A30, K2A1845, K2A2545, and K2A1850), and two other models were also included (i.e., models K2DA and K2). These six models were applied to the main NSDUH CJ data, and for each model, predicted probabilities of SMI were obtained for each respondent. Then, the cut point associated with each model was applied to the corresponding predicted probabilities to obtain the associated binary cut point estimates.

Table 10 summarizes SMI cut point estimates for each of these models overall and across the same demographic domains as those discussed in Section 4.4. For each of the six selected models, cut point estimates based on both the MHSS and NSDUH data are displayed. In almost all cases, the estimate from the NSDUH data is greater than the corresponding estimate from the MHSS data. Due to the small sample size of the MHSS data, no differences between MHSS and corresponding NSDUH estimates are statistically significant. However, due to the directional pattern of the differences (i.e., almost all NSDUH estimates are greater than corresponding MHSS estimates) and the magnitude of some of the differences, it is worth noting some of these differences. For example, for model K2A30, the MHSS estimate of SMI is 13.7% (which is reasonably close to the SCID direct estimate of 13.9%), but the NSDUH estimate is 17%, representing a substantial (but not statistically significant) increase. Estimates for the 35 or older age group also appear to be much higher in the NSDUH than in the MHSS for the models that include an age variable. For example, for model K2A30, the MHSS estimate is 15.6% (which is reasonably close to the SCID direct estimate of 16.2%), but the NSDUH estimate is 24%, representing a substantial (but not statistically significant) increase. Although it is possible that the higher estimates observed in the NSDUH data could be due to more accurate estimation due to a larger sample size, they could also be due to the cut point not being consistent with the NSDUH data. The latter can easily be checked by deriving hybrid cut points that are determined so that the SMI cut point estimate is as equal as possible to the weighted mean of the predicted probabilities in the *main NSDUH data*. Of the 23 domain-level estimates among the models in the MHSS in Table 10, 16 (69.6%) would be suppressed due to low precision according to NSDUH suppression rules.

Table 10. SMI Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH

Model		**K2A30**		**K2A1845**		**K2A2545**		**K2A1850**		**K2DA**		**K2**	
Dataset	**MHSS**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**
Estimator	**SCID**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**
Sample Size	**316**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**
Demographic Variable													
Total	13.9%	13.7%	17.0%	13.6%	15.8%	13.6%	15.0%	13.4%	15.2%	14.0%	15.0%	16.0%	14.1%
Sex													
Male	11.3%	11.1%	12.2%	11.1%	11.1%	11.1%	10.5%	10.9%	10.7%	11.4%	10.4%	12.9%	11.5%
Female	*22.7%*	*22.4%*	29.4%	*21.9%*	27.8%	*21.9%*	26.8%	*21.9%*	26.9%	*22.7%*	26.7%	*26.4%*	20.7%
Age													
18-24	7.6%	9.0%	10.8%	8.4%	7.9%	8.4%	8.1%	8.0%	7.6%	9.8%	11.8%	*17.6%*	13.6%
25-34	*16.3%*	*15.1%*	14.4%	*15.1%*	13.7%	*15.1%*	12.5%	*15.1%*	13.1%	*19.1%*	15.8%	*23.2%*	13.7%
35 or Older	*16.2%*	*15.6%*	23.6%	*15.6%*	23.4%	*15.6%*	22.1%	*15.6%*	22.5%	*13.7%*	16.8%	*11.3%*	14.7%
Race/Hispanic Origin													
Non-Hispanic White	16.4%	16.3%	22.4%	16.1%	21.3%	16.1%	20.3%	16.1%	20.4%	16.8%	19.6%	17.1%	14.8%
Non-Hispanic Black	*12.0%*	*20.2%*	10.3%	*20.2%*	8.8%	*20.2%*	8.4%	*20.2%*	8.7%	*20.2%*	8.7%	*13.7%*	13.6%
Hispanic	*2.2%*	*3.0%*	8.8%	*2.8%*	7.9%	*2.8%*	7.0%	*2.8%*	7.4%	*2.9%*	8.0%	*21.0%*	12.7%
Non-Hispanic Other	*21.8%*	*5.2%*	15.4%	*5.2%*	14.7%	*5.2%*	13.9%	*3.9%*	13.8%	*5.2%*	15.8%	*3.3%*	13.0%
Education													
Less than High School	*13.6%*	*6.3%*	15.9%	*6.1%*	14.2%	*6.1%*	13.1%	*6.1%*	13.5%	*6.3%*	13.5%	*15.6%*	17.6%
High School Graduate	13.7%	*21.0%*	14.8%	*20.6%*	13.6%	*20.6%*	12.8%	*20.3%*	12.9%	*19.4%*	13.7%	*22.5%*	14.5%
Some College	*16.4%*	*17.6%*	19.8%	*17.6%*	19.0%	*17.6%*	18.5%	*17.6%*	18.5%	*20.9%*	18.3%	*12.6%*	10.7%
College Graduate	*8.3%*	*4.2%*	22.9%	*4.2%*	22.4%	*4.2%*	21.9%	*4.2%*	22.2%	*4.2%*	16.3%	*1.8%*	8.5%
Marital Status													
Married	*17.1%*	*10.4%*	17.3%	*10.4%*	16.6%	*10.4%*	16.0%	*10.4%*	16.1%	*8.0%*	12.7%	*13.5%*	13.2%
Widowed	*1.3%*	*1.9%*	*34.9%*	*1.9%*	*29.5%*	*1.9%*	*28.9%*	*1.9%*	*29.5%*	*1.9%*	19.7%	*1.3%*	*18.3%*
Divorced or Separated	*22.2%*	*34.9%*	26.4%	*34.8%*	26.5%	*34.8%*	24.6%	*34.8%*	25.0%	*34.7%*	22.5%	*17.6%*	15.7%
Never Married	11.0%	11.6%	13.0%	11.3%	11.3%	11.3%	10.9%	11.1%	11.0%	13.6%	13.1%	18.1%	13.7%

(continued)

Table 10. SMI Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH (continued)

Model		**K2A30**		**K2A1845**		**K2A2545**		**K2A1850**		**K2DA**		**K2**	
Dataset	**MHSS**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**
Estimator	**SCID**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**
Sample Size	**316**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**
Demographic Variable													
Military Service													
Yes	*10.8%*	*12.5%*	22.3%	*12.5%*	21.5%	*12.5%*	20.1%	*12.5%*	20.6%	*12.5%*	15.3%	*17.5%*	10.5%
No	14.0%	13.8%	16.5%	13.6%	15.3%	13.6%	14.6%	13.5%	14.7%	14.1%	14.9%	15.9%	14.4%
Received Mental Health Treatment in Past Year													
Yes	*42.8%*	*46.3%*	54.5%	*45.9%*	52.3%	*45.9%*	50.3%	*45.9%*	50.9%	*43.5%*	49.4%	*47.8%*	33.4%
No	7.6%	6.7%	6.7%	6.6%	5.8%	6.6%	5.3%	6.4%	5.4%	7.7%	5.5%	9.1%	8.8%
Illicit Drug or Alcohol Abuse or Dependence in Past Year													
Yes	*15.7%*	*21.8%*	21.8%	*21.7%*	20.0%	*21.7%*	19.2%	*21.7%*	19.3%	*20.5%*	20.3%	*20.1%*	20.5%
No	*12.6%*	7.9%	14.1%	7.7%	13.3%	7.7%	12.5%	7.5%	12.7%	9.4%	11.7%	13.0%	10.2%

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. Standard errors of estimates in this table are contained in Table B.10 in Appendix B.

Predictor variables included in the following models:

K2A30: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE30
K2A1845: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1845
K2A2545: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE2545
K2A1850: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1850
K2DA: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U
K2: K6SCMON2

Datasets = CJ subsample of 2008-2012 adult NSDUH data for NSDUH estimates, and CJ subsample of 2008-2012 MHSS clinical data for MHSS estimates.
Analysis weights = ANALWT for NSDUH estimates, and MHFNLWGT for MHSS estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 11 summarizes SMI hybrid cut point estimates for each of the same six models overall and across the same demographic domains as those discussed in Section 4.4. The table also displays the original and hybrid cut points for each model. For some models, the hybrid cut point is substantially larger than the original (e.g., for model K2A30, they are 0.29126 and 0.24589, respectively), but for model K2, the two cut points are identical. For all models, the overall hybrid NSDUH estimate of SMI is much closer to the SCID direct estimate of 13.9%. However, for the four models that include an age variable, there still appears to be some bias within the two older age groups. Model K2DA (which does not include an age variable) appears to exhibit the least amount of bias within the two older age groups. This is an important result because a parsimonious model that does not include covariates related to any sociodemographic domains (e.g., age group) is preferred. A covariate related to such a domain would tend to fix the relationship between SMI and that domain. Inclusion of such a covariate would only be preferable if it provided substantial benefits to estimation, such as the substantial reduction of bias at the domain level, which was the case for the 2012 NSDUH model with respect to age groups. Therefore, model K2DA was selected to proceed to the next step of assessing SMI estimates within subpopulations of the CJ population itself. Of the 23 domain-level SCID direct estimates in the MHSS in Table 11, 16 (69.6%) would be suppressed due to low precision according to NSDUH suppression rules.

Table 11. SMI Hybrid Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH

Model		**K2A30**	**K2A1845**	**K2A2545**	**K2A1850**	**K2DA**	**K2**
Original Cut Point		**0.24589**	**0.26490**	**0.28900**	**0.27548**	**0.25776**	**0.21262**
Hybrid Cut Point		**0.29126**	**0.30852**	**0.30971**	**0.29772**	**0.27879**	**0.21262**
Dataset	**MHSS**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**
Estimator	**SCID**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**
Sample Size	**316**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**
Demographic Variable							
Total	13.9%	14.1%	13.9%	13.9%	13.9%	14.3%	14.1%
Sex							
Male	11.3%	9.7%	9.7%	9.6%	9.7%	9.8%	11.5%
Female	*22.7%*	25.4%	24.7%	24.9%	24.6%	25.9%	20.7%
Age							
18-24	7.6%	9.1%	5.8%	6.3%	6.6%	11.3%	13.6%
25-34	*16.3%*	13.0%	12.5%	11.8%	12.5%	15.1%	13.7%
35 or Older	*16.2%*	18.8%	21.1%	21.2%	20.5%	15.9%	14.7%
Race/Hispanic Origin							
Non-Hispanic White	16.4%	18.8%	18.6%	18.8%	18.7%	18.8%	14.8%
Non-Hispanic Black	*12.0%*	8.3%	8.0%	8.1%	7.7%	8.3%	13.6%
Hispanic	*2.2%*	6.9%	6.7%	6.3%	6.7%	7.8%	12.7%
Non-Hispanic Other	*21.8%*	13.5%	12.9%	12.9%	12.9%	13.1%	13.0%
Education							
Less than High School	*13.6%*	12.6%	12.3%	12.2%	12.2%	12.8%	17.6%
High School Graduate	13.7%	12.3%	11.4%	11.4%	11.4%	13.1%	14.5%
Some College	*16.4%*	17.2%	17.3%	17.4%	17.2%	17.3%	10.7%
College Graduate	*8.3%*	19.2%	21.1%	21.1%	21.2%	15.9%	8.5%
Marital Status							
Married	*17.1%*	13.4%	15.1%	14.8%	14.4%	11.8%	13.2%
Widowed	*1.3%*	*26.8%*	*25.8%*	*27.2%*	*28.6%*	19.7%	*18.3%*
Divorced or Separated	*22.2%*	22.7%	24.0%	24.2%	24.1%	21.4%	15.7%
Never Married	11.0%	11.0%	9.5%	9.5%	9.6%	12.6%	13.7%

(continued)

Table 11. SMI Hybrid Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH (continued)

Model		**K2A30**	**K2A1845**	**K2A2545**	**K2A1850**	**K2DA**	**K2**
Original Cut Point		**0.24589**	**0.26490**	**0.28900**	**0.27548**	**0.25776**	**0.21262**
Hybrid Cut Point		**0.29126**	**0.30852**	**0.30971**	**0.29772**	**0.27879**	**0.21262**
Dataset	**MHSS**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**
Estimator	**SCID**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**
Sample Size	**316**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**
Demographic Variable							
Military Service							
Yes	*10.8%*	18.7%	19.4%	18.9%	19.2%	14.9%	10.5%
No	14.0%	13.7%	13.4%	13.4%	13.4%	14.2%	14.4%
Received Mental Health Treatment in Past Year							
Yes	*42.8%*	47.9%	47.3%	47.7%	46.8%	47.8%	33.4%
No	7.6%	4.9%	4.7%	4.6%	4.8%	5.1%	8.8%
Illicit Drug or Alcohol Abuse or Dependence in Past Year							
Yes	*15.7%*	18.5%	17.6%	17.6%	17.9%	19.3%	20.5%
No	*12.6%*	11.5%	11.6%	11.6%	11.4%	11.2%	10.2%

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. Standard errors of estimates in this table are contained in Table B.11 in Appendix B. Original cut point derived to yield (nearly) unbiased cut point estimator in MHSS, and hybrid cut point derived to yield (nearly) unbiased cut point estimator in NSDUH.

Predictor variables included in the following models:

K2A30: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE30
K2A1845: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1845
K2A2545: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE2545
K2A1850: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1850
K2DA: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U
K2: K6SCMON2

Datasets = CJ subsample of 2008-2012 adult NSDUH data for NSDUH estimates, and CJ subsample of 2008-2012 MHSS clinical data for MHSS estimates.
Analysis weights = ANALWT for NSDUH estimates, and MHFNLWGT for MHSS estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 12 summarizes SMI hybrid cut point estimates for model K2DA within several subpopulations of the CJ population, overall and across the same demographic domains as those discussed in Section 4.4. The CJ subpopulations include (1) entire CJ population, (2) parolees, (3) probationers, and (4) parolees or probationers. The overall sample sizes of these subpopulations are also displayed in the table. For example, within the MHSS, the samples sizes for the subpopulations are: parolees (n = 56), probationers (n = 186), and parolees or probationers (n = 195). In the MHSS, the sample size for parolees is too small to analyze this subpopulation separately, and because most parolees also reported being on probation, the combined parolee and probationer subpopulation was considered along with the entire CJ population. Within the MHSS, the SCID direct estimates differ markedly between the entire CJ population and the parolee or probationer subpopulation. However, within the NSDUH, the hybrid cut point estimates are all fairly similar within the two subpopulations. The stability in the NSDUH hybrid cut point estimates between the two subpopulations suggests that the variation in the MHSS SCID direct estimates between the same two subpopulations is likely due to the small sample sizes of these subpopulations in the MHSS. This further suggests that estimates based on the parolee or probationer subpopulation (which is assumed to be more similar to the prisoner population than the entire CJ population that also includes arrestees) appear to apply similarly to the entire CJ population. Note that the hybrid cut points are fairly similar between the two subpopulations (0.27879 and 027659, respectively). Of the 23 domain-level SCID direct estimates for the entire CJ population in the MHSS in Table 12, 16 (69.6%) would be suppressed due to low precision according to NSDUH suppression rules. All other total and domain-level SCID direct estimates for subgroups within the CJ population in the MHSS in Table 12 would be suppressed.

The next step was to extrapolate the K2DA model, cut point from the CJ subpopulation in the MHSS, and two NSDUH hybrid cut points from the two subpopulations (entire NSDUH CJ population and the NSDUH parolees and probationers) to the prisoner data from the 2016 SPI.

Table 12. SMI Hybrid Cut Point Estimates of Model K2DA for Criminal Justice Subpopulations, 2008-2012 NSDUH

Criminal Justice Subpopulation	All		Parolees		Probationers		Parolees or Probationers	
Dataset	MHSS	NSDUH	MHSS	NSDUH	MHSS	NSDUH	MHSS	NSDUH
Estimator	SCID	Hybrid	SCID	Hybrid	SCID	Hybrid	SCID	Hybrid
Hybrid Cut Point		0.27879		0.25525		0.27879		0.27659
Sample Size	316	15,400	56	2,300	186	8,300	195	8,800
Demographic Variable								
Total	13.9%	14.3%	*11.6%*	13.2%	*16.2%*	14.6%	*16.0%*	14.4%
Sex								
Male	11.3%	9.8%	*9.2%*	9.7%	*12.0%*	10.1%	*13.0%*	9.8%
Female	*22.7%*	25.9%	*16.5%*	25.8%	*32.2%*	25.6%	*25.6%*	25.8%
Age								
18-24	7.6%	11.3%	*0.6%*	13.3%	*10.9%*	11.5%	*10.4%*	11.6%
25-34	*16.3%*	15.1%	*27.8%*	13.9%	*16.6%*	14.9%	*16.4%*	14.4%
35 or Older	*16.2%*	15.9%	*11.0%*	12.8%	*18.8%*	16.8%	*18.7%*	16.4%
Race/Hispanic Origin								
Non-Hispanic White	16.4%	18.8%	*10.9%*	18.1%	*25.6%*	19.5%	*23.1%*	19.3%
Non-Hispanic Black	*12.0%*	8.3%	*24.2%*	10.5%	*6.2%*	9.0%	*16.1%*	8.3%
Hispanic	*2.2%*	7.8%	*0.1%*	7.0%	*2.7%*	5.9%	*2.5%*	6.3%
Non-Hispanic Other	*21.8%*	13.1%	*77.2%*	*14.4%*	*6.4%*	14.0%	*6.4%*	14.9%
Education								
Less than High School	*13.6%*	12.8%	*19.4%*	13.0%	*12.5%*	12.4%	*15.8%*	12.2%
High School Graduate	13.7%	13.1%	*4.7%*	13.6%	*14.8%*	13.9%	*12.4%*	13.6%
Some College	*16.4%*	17.3%	*9.6%*	13.7%	*25.5%*	18.6%	*24.4%*	18.6%
College Graduate	*8.3%*	15.9%	*65.8%*	*12.1%*	*10.8%*	13.2%	*10.7%*	13.0%
Marital Status								
Married	*17.1%*	11.8%	*3.4%*	9.6%	*15.2%*	11.3%	*12.5%*	11.4%
Widowed	*1.3%*	19.7%	*	*20.6%*	*	*25.5%*	*	*24.6%*
Divorced or Separated	*22.2%*	21.4%	*24.8%*	20.2%	*24.9%*	23.5%	*30.4%*	22.6%
Never Married	11.0%	12.6%	*11.9%*	11.9%	*14.6%*	12.6%	*14.3%*	12.3%

(continued)

Table 12. SMI Hybrid Cut Point Estimates of Model K2DA for Criminal Justice Subpopulations, 2008-2012 NSDUH (continued)

Criminal Justice Subpopulation	**All**		**Parolees**		**Probationers**		**Parolees or Probationers**	
Dataset	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**
Estimator	**SCID**	**Hybrid**	**SCID**	**Hybrid**	**SCID**	**Hybrid**	**SCID**	**Hybrid**
Hybrid Cut Point		**0.27879**		**0.25525**		**0.27879**		**0.27659**
Sample Size	**316**	**15,400**	**56**	**2,300**	**186**	**8,300**	**195**	**8,800**
Demographic Variable								
Military Service								
Yes	*10.8%*	14.9%	*	*6.6%*	*19.4%*	12.6%	*19.4%*	11.7%
No	14.0%	14.2%	*12.3%*	13.8%	*16.0%*	14.8%	*15.9%*	14.6%
Received Mental Health Treatment in Past Year								
Yes	*42.8%*	47.8%	*61.6%*	43.5%	*51.3%*	46.9%	*52.5%*	46.6%
No	7.6%	5.1%	*4.4%*	5.2%	5.7%	4.8%	5.2%	4.9%
Illicit Drug or Alcohol Abuse or Dependence in Past Year								
Yes	*15.7%*	19.3%	*17.0%*	17.9%	*20.3%*	19.1%	*20.3%*	19.0%
No	*12.6%*	11.2%	*9.4%*	11.1%	*12.8%*	12.2%	*13.0%*	11.8%

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

*Indicates actual suppression due to disclosure risk from small denominators.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. Standard errors of estimates in this table are contained in Table B.12 in Appendix B. Hybrid cut point derived to yield (nearly) unbiased cut point estimator in NSDUH.

Predictor variables included in the following model:

K2DA: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U

Datasets = CJ subsample of 2008-2012 adult NSDUH data for NSDUH estimates, and CJ subsample of 2008-2012 MHSS clinical data for MHSS estimates.
Analysis weights = ANALWT for NSDUH estimates, and MHFNLWGT for MHSS estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

4.7 Initial Application and Assessment of Models in 2016 SPI

In the final step of predicting SMI among the prisoner data of the 2016 SPI, the SMI prediction model K2DA, developed within the CJ subsample of the MHSS, was selected for extrapolation to the 2016 SPI. In addition, three cut points associated with model K2DA also were used on the 2016 SPI data: the original cut point determined from the CJ subsample of the MHSS (i.e., 0.25776), and the two NSDUH hybrid cut points determined from the entire CJ population (0.27879) and the CJ subpopulation of parolees or probationers (0.27659).

Three SMI cut point estimates were obtained by applying model K2DA to the 2016 SPI data for each of the three cut points. In each of the three cases, the cut point estimate was about 3 to 4 percentage points larger than the mean of the predicted probabilities of SMI in the SPI data. Therefore, a hybrid cut point was determined from within the SPI data itself (i.e., 0.37446), and this resulted in a fourth cut point estimate of SMI obtained from the 2016 SPI data. A hybrid cut point within the SPI data was obtained by selecting the cut point so that the SMI cut point estimate was as equal as possible to the weighted mean of the predicted probabilities of SMI derived from the SPI data.

In addition, cut point estimates were obtained for several domains within the 2016 SPI data. Although there is some overlap in domain-level information between the 2008-2012 NSDUH and the 2016 SPI, some domains exist in the 2008-2012 NSDUH that do not exist in the 2016 SPI, and vice versa.

For domains where there was overlap between the two surveys, domain-level information in the 2016 SPI was recoded to conform with the 2008-2012 NSDUH domains described in Section 4.5. These include the following domains:

- sex (male, female)
- age (18 to 24, 25 to 34, 35 or older)
- race/Hispanic origin (non-Hispanic white, non-Hispanic black, Hispanic, non-Hispanic other)
- marital status (married, widowed, divorced or separated, never married)
- military service ("yes," "no").

Two other domains included in the 2016 SPI (but not in the 2008-2012 NSDUH) of particular interest are:

- jurisdiction (federal, state)
- offense (violent, property, drug, public order, other).

Four sets of cut point estimates of SMI (overall and at the domain level) corresponding to the four cut points described above are displayed in Table 13. For comparative purposes, SMI estimates from both the MHSS and NSDUH are shown in the same table. This table includes only the domains associated with the 2016 SPI.

Table 13. SMI Estimates Based on Model K2DA for Different Data, Cut Points, and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI

Dataset	2008-2012 MHSS			2008-2012 NSDUH			2016 SPI			
Subpopulation	**Criminal Justice**		**Parolees or Probationers**	**Criminal Justice**		**Parolees or Probationers**	**All Prisoners**			
Estimator	**SCID**	**MHSS Cut Point**	**SCID**	**MHSS Cut Point**	**CJ Hybrid Cut Point**	**PP Hybrid Cut Point**	**MHSS Cut Point**	**CJ Hybrid Cut Point**	**PP Hybrid Cut Point**	**SPI Hybrid Cut point**
Cut Point		**0.25776**		**0.25776**	**0.27879**	**0.27659**	**0.25776**	**0.27879**	**0.27659**	**0.37446**
Sample Size	**316**	**316**	**195**	**15,400**	**15,400**	**8,800**	**24,848**	**24,848**	**24,848**	**24,848**
Demographic Variable										
Total	13.9%	14.0%	*16.0%*	15.0%	14.3%	14.4%	20.5%	19.6%	19.7%	16.7%
Sex										
Male	11.3%	11.4%	*13.0%*	10.4%	9.8%	9.8%	18.9%	18.1%	18.2%	15.2%
Female	*22.7%*	*22.7%*	*25.6%*	26.7%	25.9%	25.8%	41.4%	40.2%	40.4%	36.7%
Age										
18-24	7.6%	9.8%	*10.4%*	11.8%	11.3%	11.6%	19.0%	18.2%	18.2%	15.4%
25-34	*16.3%*	*19.1%*	*16.4%*	15.8%	15.1%	14.4%	21.6%	20.6%	20.8%	17.4%
35 or Older	*16.2%*	*13.7%*	*18.7%*	16.8%	15.9%	16.4%	20.1%	19.3%	19.4%	16.6%
Race/Hispanic Origin										
Non-Hispanic White	16.4%	16.8%	*23.1%*	19.6%	18.8%	19.3%	26.9%	26.1%	26.2%	22.9%
Non-Hispanic Black	*12.0%*	*20.2%*	*16.1%*	8.7%	8.3%	8.3%	14.6%	13.8%	13.8%	10.7%
Hispanic	*2.2%*	*2.9%*	*2.5%*	8.0%	7.8%	6.3%	16.7%	16.0%	16.1%	13.6%
Non-Hispanic Other	*21.8%*	*5.2%*	*6.4%*	15.8%	13.1%	14.9%	27.2%	26.0%	26.1%	22.7%
Marital Status										
Married	*17.1%*	*8.0%*	*12.5%*	12.7%	11.8%	11.4%	19.4%	18.7%	18.9%	16.1%
Widowed	*1.3%*	*1.9%*	*	19.7%	19.7%	*24.6%*	19.7%	19.3%	19.3%	18.1%
Divorced or Separated	*22.2%*	*34.7%*	*30.4%*	22.5%	21.4%	22.6%	23.7%	22.8%	22.8%	20.0%
Never Married	11.0%	13.6%	*14.3%*	13.1%	12.6%	12.3%	19.5%	18.6%	18.7%	15.4%
Military Service										
Yes	*10.8%*	*12.5%*	*19.4%*	15.3%	14.9%	11.7%	22.0%	21.2%	21.4%	18.1%
No	14.0%	14.1%	*15.9%*	14.9%	14.2%	14.6%	20.4%	19.5%	19.6%	16.6%

(continued)

Table 13. SMI Estimates Based on Model K2DA for Different Data, Cut Points, and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI (continued)

Dataset	2008-2012 MHSS			2008-2012 NSDUH			2016 SPI			
Subpopulation	Criminal Justice		Parolees or Probationers	Criminal Justice		Parolees or Probationers	All Prisoners			
Estimator	SCID	MHSS Cut Point	SCID	MHSS Cut Point	CJ Hybrid Cut Point	PP Hybrid Cut Point	MHSS Cut Point	CJ Hybrid Cut Point	PP Hybrid Cut Point	SPI Hybrid Cut point
Cut Point		0.25776		0.25776	0.27879	0.27659	0.25776	0.27879	0.27659	0.37446
Sample Size	316	316	195	15,400	15,400	8,800	24,848	24,848	24,848	24,848
Demographic Variable										
Jurisdiction										
Federal	~	~	~	~	~	~	9.8%	9.0%	9.1%	7.6%
State	~	~	~	~	~	~	22.0%	21.1%	21.2%	18.0%
Offense										
Violent	~	~	~	~	~	~	20.8%	19.8%	19.9%	16.7%
Property	~	~	~	~	~	~	24.4%	23.7%	23.8%	20.4%
Drug	~	~	~	~	~	~	18.3%	17.6%	17.7%	15.1%
Public Order	~	~	~	~	~	~	23.8%	23.6%	23.6%	21.6%
Other	~	~	~	~	~	~	:	:	:	:

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; PP = parolees or probationers; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness; SPI = Survey of Prison Inmates.

55

:Not calculated.

~Not applicable.

*Indicates actual suppression due to disclosure risk from small denominators.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. Standard errors of estimates in this table are contained in Table B.13 in Appendix B.

Datasets = 2008-2012 MHSS clinical data for MHSS estimates, 2008-2012 adult NSDUH data for NSDUH estimates, and 2016 SPI data for SPI estimates.
Analysis weights = MHFNLWGT for MHSS estimates, ANALWT for NSDUH estimates, and WT_FINAL for SPI estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012; and Bureau of Justice Statistics, Survey of Prison Inmates, 2016.

In Table 13, the first three columns of SMI estimates are associated with the 2008-2012 MHSS sample. The first column is associated with directly computed SCID estimates and the second with cut point estimates within the entire CJ population. These two sets of estimates closely agree overall (13.9%) and across most domains (14.0%). The third column is associated with SCID direct estimates within the CJ subpopulation of parolees or probationers (excluding those only reporting arrest). These estimates are somewhat higher (overall estimate is 16.0%). No cut point estimates for subpopulations within the CJ population (e.g., parolees or probationers) were computed due to the small sample sizes of those subpopulations (e.g., n = 195 for parolees or probationers).

The next three columns of SMI estimates in Table 13 are associated with the 2008-2012 NSDUH. The first two columns are associated with cut point estimates for the entire CJ population. The first set of estimates are associated with the original MHSS cut point, and the second set with the NSDUH hybrid cut point determined from the entire CJ population. The third column is associated with cut point estimates within the CJ subpopulation of parolees or probationers, using the NSDUH hybrid cut point for that subpopulation. The three sets of estimates agree reasonably well (overall estimates are 14.3% to 15.0%), and these estimates do not differ greatly from those of the first two MHSS columns.

The final four columns of SMI estimates in Table 13 are associated with the 2016 SPI. These estimates are based on model K2DA, and each of the four cut points are applied to the 2016 SPI data. The first three columns are associated with MHSS or NSDUH hybrid cut points, and the estimates agree reasonably well (overall estimates are 19.6% to 20.5%), but these are substantially larger than the estimates associated with the MHSS or NSDUH data. The final column of estimates is associated with the SPI cut point, and these estimates are somewhat smaller than those associated with the MHSS or NSDUH cut points (overall estimate is 16.7%). All four columns of SPI estimates in Table 13 show that females have higher rates of SMI than males, and non-Hispanic whites have higher rates than non-Hispanic blacks and Hispanics. These results are consistent with those among the CJ population in the NSDUH and with results from three BJS reports, all of which used different methodologies. Bronson and Berzofsky (2017) showed higher rates among females and higher rates among non-Hispanic whites compared to non-Hispanic blacks and Hispanics. James and Glaze (2006) and Ditton (1999) had similar results. All four columns of SPI estimates in Table 13 show similar rates of SMI among the three age groups. This is consistent with Bronson and Berzofsky (2017), which shows no differences in rates of SPD by age among prisoners (state and federal). The SPI estimates show lower rates among federal prisoners, which is consistent with prior BJS research, such as James and Glaze (2006) and Ditton (1999), and in the 2016 SPI, federal prisoners have lower rates of SPD than state prisoners. The characteristics of the federal and state prison populations are different. The proportion of females is slightly smaller in federal (6.2%) than state (7.1%) prisons. There is a smaller proportion of non-Hispanic white persons in federal (21.0%) than state (31.9%) prisons, and a slightly smaller proportion of non-Hispanic black persons in federal (32.2%) compared to state (33.8%) prisons. Federal prisoners (mean age of 40.4) are slightly older than state prisoners (age 39.0), with 32.9% of federal prisoners age 34 or younger, compared to 42.4% of state prisoners, and 4.7% of federal prisoners age 24 or younger, compared to 10.3% of state prisoners. Almost half (47.7%) of federal prisoners are incarcerated for a drug offense, compared to more than half (56.2%) of state prisoners incarcerated for a violent offense. Bronson and Berzofsky (2017) shows higher rates of SPD among prisoners sentenced for a violent offense compared to those sentenced for a drug offense.

4.8 SMI Estimation in 2016 SPI: Considerations

There are several reasons why the overall estimates of SMI might range from 13.9% to 20.5% across all 10 sets of estimates in Table 13. First, the *sample size* of the CJ subsample of the 2008-2012 MHSS from which the K2DA model was developed and the MHSS cut point determined is small (n = 316). Figures 1 and 2 in Section 4.1.5 illustrate the impact of sample size on the distribution of predicted probabilities from a model, and on the precision of the associated cut point. In addition, the predicted probabilities and estimates of the model parameters will be less precise due to the small sample size.

In Figure 1, a CDF plot of predicted probabilities from the 2012 NSDUH model obtained from the 2008A-2012 MHSS (n = 4,912) is smooth, thus allowing for a precisely determined cut point. This is confirmed by the fact that the same model extrapolated to the full 2008A-2012 adult NSDUH provides an almost identically overlapping CDF of predicted probabilities, and that the hybrid cut point determined from this larger dataset is almost identical to the original cut point determined from the MHSS subsample.

In comparison, Figure 2 shows a CDF plot of predicted probabilities from SMI model K2DA obtained from the CJ subsample of the 2008-2012 MHSS (n = 316), which is not smooth, and this presents difficulties in precisely determining an appropriate cut point. This is confirmed by the fact that the same model extrapolated to the CJ subsample of the 2008-2012 adult NSDUH does not provide CDFs that consistently overlap, and that the hybrid cut point determined from this larger dataset is somewhat different from the original cut point determined from the subsample.

In addition, the further subsetting of the CJ subsample of the 2008-2012 MHSS to the CJ subpopulation of parolees or probationers (n = 195) exacerbates the problems associated with small sample sizes discussed above.

To summarize the issue of small sample sizes with respect to Table 13: (1) the SCID direct estimates are based on small sample sizes (n = 316 for the CJ estimates, and n = 195 for the parolee and probationer estimates); (2) all remaining cut point estimates in the table are based on SMI model K2DA (developed from a dataset where n = 316); and (3) the MHSS cut point was determined from a dataset where n = 316. In spite of this, with the exception of the parolee and probationer subpopulation estimates, MHSS and NSDUH estimates of SMI agree reasonably well.

Second, *extrapolation* of a model (and cut point) from a subsample of a larger sample to the larger sample itself appears to be reasonable if the size of the subsample is sufficiently large (e.g., Figure 1 suggests this was true for the 2008-2012 MHSS). However, these SPI estimates are based on extrapolation of a model (and cut point) from a small subsample of a clinical sample from one survey (i.e., CJ subsample of 2008-2012 MHSS) to an entirely different survey (i.e., the 2016 SPI). The two surveys possess different methodological characteristics (e.g., sampling frame, data collection methods). In particular, the sampling frames of the 2008-2012 NSDUH (from which the 2008-2012 MHSS was subsampled) and the 2016 SPI have been designed to have almost no overlap. The 2008-2012 NSDUH specifically excluded prisoners from the sampling frame, whereas the 2016 SPI included only prisoners in its sampling frame (but it is still possible that someone from the CJ population in 2008 to 2012 may have been in prison in 2016). In addition, the surveys collect data in different settings (NSDUH typically in respondents' homes; SPI in prisons) using different modes [Audio Computer-Assisted Self-Interviewing (ACASI) for NSDUH and CAPI for SPI], which are likely to result in different reporting behavior. These survey design differences could account for some of the reason why the SPI hybrid cut point (i.e., 0.37446) is substantially different from the range of the three cut points based on the MHSS or NSDUH (i.e., 0.25776 to 0.27879).

Third, there is *limited common information* between the 2016 SPI and the NSDUH available for the development of SMI prediction models. The 2016 SPI did not collect information on the past year K6 scale, the WHODAS impairment measure, past year suicidality, or past year MDE. All of these variables were used as predictor variables for the 2012 NSDUH prediction model in the MHSS. The 2016 SPI did collect information on the past month K6 scale, lifetime depression, and lifetime anxiety, all of which were used in the SMI model K2DA. However, these SPI mental health indicators might not predict SMI as accurately as the variables available in NSDUH.

Finally, the *populations* defined by criminal justice status (even if subsetted to parolees or probationers only) in the 2008-2012 NSDUH and actual prisoners in the 2016 SPI are likely to differ at least to some extent. Some differences may be due simply to the different time frames of the two data sources: 2008 to 2012 for NSDUH and 2016 for SPI. But the major difference between these two populations is that the first is currently not incarcerated and the second is. Of course, parolees and possibly some probationers and arrestees also were formerly incarcerated, which means that difference likely will not be as large as between prisoners and the non-CJ population because the surveys cover many of the same populations but interview them at different times (and living situations and modes) during their lives. Therefore, the higher rates of SMI estimated in the 2016 SPI may be a consequence of prisoners experiencing higher levels of psychological distress, depression, or anxiety. However, there is no way to test this assumption with the current information available, but it does appear to be consistent with the findings of Bronson and Berzofsky (2017) that rates of psychological distress, depression, or anxiety are higher in the prisoner population than among the general population.

4.9 Selection of Final SMI Model in the 2016 SPI

Analyses in this chapter have shown that psychological distress as measured by the past month K6 score and diagnosis of depression and anxiety during a person's lifetime are associated with SMI among the CJ population of the 2008-2012 MHSS and NSDUH. Analyses have also shown that the SMI model K2DA (which includes terms representing alternative past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety) provides cut point estimates within the CJ subsample of the 2008-2012 MHSS that are similar to gold-standard SCID direct estimates of SMI overall and across most domains. The same model extrapolated to the CJ subsample of the adult 2008-2012 NSDUH provides similar estimates (overall and across most domains) after new hybrid cut points were determined. The hybrid cut points ensure that the resulting cut point estimates are consistent with the predicted probabilities of the model extrapolated to the larger dataset; that is, the hybrid cut point is determined so that the cut point estimator is as close to the mean of the predicted probabilities as possible. The need to determine hybrid cut points in the larger dataset is because the CDF of predicted probabilities of SMI model K2DA is not smooth in the small CJ subsample of the 2008-2012 MHSS. Therefore, the original cut point determined in that subsample is likely to be imprecise (see Figure 2).

If the association between past month K6 score and lifetime self-reported diagnosed depression and lifetime self-reported diagnosed anxiety are associated with SMI among the prisoner population similarly to the observed association among the CJ population of the 2008-2012 MHSS and NSDUH, then SMI model K2DA is likely to provide reasonable predicted probabilities of SMI among the prisoner population. As noted in Section 4.8, the higher predicted probabilities observed among the prisoner population may be due to the fact that this population differs somewhat from the CJ population not currently incarcerated. Their true estimates while in prison may indeed be higher. Also, the three cut points extrapolated from the MHSS and NSDUH provide cut point estimates of SMI in the 2016 SPI data that are inconsistent with the predicted probabilities within the 2016 SPI data. (This is likely due to having to extrapolate the cut points from one survey to a second completely different survey.) Therefore, cut point estimates of SMI among prisoners in the 2016 SPI that are consistent with the predicted probabilities of the model used to derive the cut point estimator are most likely to reflect the prevalence of SMI among those prisoners.

Therefore, the SMI prediction model for the 2016 SPI can be expressed as—

$$\begin{aligned}\text{logit}(\hat{\pi}) \equiv \log\left[\hat{\pi}/(1-\hat{\pi})\right] = -2.7108537 + 0.1093374 * X_{km} + 1.5438971 * X_{dep} \\ +0.7763546 * X_{anx},\end{aligned}$$

where $\hat{\pi}$ is the estimated probability that an adult had SMI, and the predictor variables X_{km}, X_{dep}, and X_{anx} are defined as—

> X_{km} = *Alternative Past Month K6 Score:* past month K6 score of less than 8 recoded as 0; past year K6 score of 8 to 24 recoded as 1 to 17 [i.e., **K6SCMON2**].
>
> X_{dep} = *Lifetime Depression*: coded as 1 if "yes"; coded as 0 otherwise [i.e., **DEPLIF**].
>
> X_{anx} = *Lifetime anxiety*: coded as 1 if "yes"; coded as 0 otherwise [i.e., **ANXLIF**].

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

$$\text{SMIPP} = 1 \Big/ \left(1 + \exp\left[-\begin{pmatrix}-2.7108537 + 0.1093374 * X_{km} + 1.5438971 * X_{dep} \\ +0.7763546 * X_{anx}\end{pmatrix}\right]\right).$$

If SMIPP is greater than or equal to 0.3744640883 (the SMI hybrid cut point from the 2016 SPI data), then the respondent is predicted as having past year SMI. Otherwise, the respondent is predicted as not having past year SMI.

5. Estimating AMI Among Prisoners

5.1 General Estimation Approach and Assumptions

Although SMI estimation is the primary focus of this report, an important and secondarily focus is on AMI estimation. Section 2.5 explains that AMI is a composite measure that includes all three levels of impairment due to having a mental illness (i.e., mild, moderate, and serious). That is, AMI is the presence of mental illness regardless of the level of impairment caused by the mental illness.

Two general methods were examined for estimating AMI. One method assumes that the predicted probabilities of SMI resulting from the selected SMI prediction model constitute a range of values for different levels of mental illness. For example, the SMI cut point estimator is based on the assumption that predicted probabilities at or more than a specified cut point is an indicator of SMI. A reasonable generalization assumes that predicted probabilities at or above a different threshold could be associated with a different level of mental illness (e.g., AMI). The MHSS used this approach (CBHSQ, 2014a). (See Chapter 2 for details of the separate cut point used to create the AMI cut point estimator in the MHSS.) The advantage of this approach is its simplicity because no new AMI prediction model needs to be developed (i.e., the SMI prediction model is used), and the AMI estimates are consistent with SMI estimates (i.e., a respondent predicted as having SMI will also be predicted as having AMI). One potential disadvantage of this approach is that the SMI prediction model is developed to predict SMI specifically and may not be as effective at identifying cases that have other levels of mental illness. Also, a second disadvantage of this approach is that because the mean of the predicted probabilities of the SMI prediction model will result in an estimate of SMI (i.e., not AMI), hybrid cut points for AMI cannot be determined.

A second method relies on the development of a completely separate AMI prediction model. A cut point is then identified for this model so that predicted probabilities at or above this threshold will be associated with the presence of AMI. Besides the need to develop a separate AMI prediction model (whose predictor variables and beta estimates might be quite different from those of the SMI prediction model), the major disadvantage of this approach is that estimates of AMI are not guaranteed to be consistent with estimates of SMI. For example, a respondent may be predicted as having SMI by the SMI model and cut point but be predicted as *not* having AMI by the AMI model and cut point. Certain methods enforce consistency between SMI and AMI estimates. For example, one approach would be to identify all respondents who were predicted as having SMI, and then assign a new set of AMI predicted probabilities all equal to 1. For these respondents, the new set of predicted probabilities (all equal to 1) are guaranteed to equal or exceed any AMI cut point, which would result in those respondents being predicted as having AMI (thus consistent with being predicted as having SMI). This approach would introduce complications into the estimation process because those respondents would then have two sets of predicted probabilities (i.e., the original predicted probabilities and the new ones assigned the value of 1).

Estimates and ROC statistics using both of these methods were examined to determine the optimal way to produce estimates of AMI. For each method, the separate AMI cut point (or separate AMI model and cut point) was developed within the same CJ subsample of the 2008-2012 MHSS (n = 316). Where applicable, the same model selection approach and the same methodology to determine cut points were used. See Chapter 4 for details of the general approach and assumptions that were made for estimating SMI.

In addition, the same limited set of predictor variables that were available for SMI estimation were available for AMI estimation (i.e., they had to conform with what was available in the 2016 SPI data). These variables included (some version of) past month K6 score, lifetime self-reported diagnosed anxiety, and lifetime self-reported diagnosed depression (see Section 3.4 for descriptions of these variables). Initial crosstabulations demonstrated strong associations between each of the three mental health variables and two measures of AMI.

Table 14 compares predicted AMI estimates based on the 2012 NSDUH prediction model among the following subsamples of the 2008-2012 adult NSDUH: (1) all adults; (2) those on probation or parole; (3) arrestees, but not on probation or parole; (4) arrestees, on probation or parole (i.e., our CJ subsample); and (5) all adults excluding the CJ subsample. In Table 14, AMI estimates are also broken out by several mental health indicators based on past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety.

Table 14 indicates that an estimated 32.0% of the CJ population has AMI, which is higher than the 17.5% estimated for the non-CJ population. Within the CJ population, 32.8% is estimated to have AMI for parolees and probationers and 30.8% for the arrestee (but not on probation or parole) population. This pattern of results appears to hold across the mental health indicators shown in the table.

Among all the subsamples presented in Table 14, AMI prevalence estimates increase as past month K6 score increases and in the presence of lifetime depression and lifetime anxiety, suggesting that these mental health indicators would be good predictors of AMI in regression models.

Table 14. Predicted AMI, by Mental Health Indicator and Criminal Justice Status Among Adults 18 or Older, 2008-2012 NSDUH

	Criminal Justice Status														
	All Adults			On Probation or Parole†			Arrestee, Not on Probation or Parole			Arrestee, on Probation or Parole‡			Not Arrestee, Not on Probation or Parole		
Mental Health Indicator	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**
Total	18.1	%	0.13	32.8	%	0.89	30.8	%	1.03	32.0	%	0.69	17.5	**	0.14
Past Month K6 Score															
< 8	9.0	%	0.11	14.2	%	0.90	10.7	*	0.85	12.8	%	0.64	8.8	**	0.11
8-12	56.7	%	0.48	56.0	%	1.99	51.9	%	2.68	54.2	%	1.64	56.9	%	0.50
> 12	89.6	%	0.37	88.0	%	1.37	86.2	%	2.39	87.2	%	1.27	89.9	**	0.38
Lifetime Depression															
No	12.6	%	0.13	24.5	%	0.92	23.1	%	1.00	23.9	%	0.67	12.1	**	0.13
Yes	56.2	%	0.46	70.3	%	2.05	71.8	%	2.65	70.9	%	1.62	55.3	**	0.47
Lifetime Anxiety															
No	14.7	%	0.13	26.1	%	0.92	24.9	%	1.03	25.6	%	0.68	14.3	**	0.13
Yes	55.2	%	0.57	75.0	%	1.98	71.4	%	2.46	73.6	%	1.55	53.8	**	0.59
Past Month K6 Score, Lifetime Depression, and Lifetime Anxiety															
< 8, No Lifetime Depression or Lifetime Anxiety	6.4	%	0.10	10.7	%	0.81	8.3	*	0.83	9.7	%	0.58	6.3	**	0.10
> 12, with Lifetime Depression and Lifetime Anxiety	97.7	%	0.25	98.5	%	0.61	98.1	%	0.66	98.4	%	0.46	97.6	%	0.28
Sample Size	229,600			8,800			6,600			15,400			213,000		

AMI = any mental illness; K6 = Kessler 6-item distress scale; NSDUH = National Survey on Drug Use and Health; SE = standard error.

*Difference between this estimate and estimate from comparison group † is statistically significant at the 0.05 level.

**Difference between this estimate and estimate from comparison group ‡ is statistically significant at the 0.05 level.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules.

Dataset = 2008-2012 adult NSDUH data.
Analysis weight = ANALWT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table 15 compares gold-standard AMI direct estimates among the same subsamples of the 2008-2012 MHSS. The table also breaks out AMI estimates by several mental health indicators based on past month K6 score, lifetime depression, lifetime anxiety, and GAF score.

Table 15 indicates that an estimated 28.2% of the CJ population has AMI, which is higher than the 17.4% estimated for the non-CJ population. Within the CJ population, 36.8% is estimated to have AMI for parolees and probationers and 22.5% for the arrestee (but not on probation or parole) population. However, although these estimates appear to be substantially different, both of these estimates have large standard errors (8.16 for parolees and probationers and 5.40 for arrestees) and would have been suppressed according to the NSDUH precision-based suppression rules due to the small sample sizes. This pattern of results appears to hold across the mental health indicators shown in the table. Almost all estimates among the various CJ subpopulations represented in Table 15 would be suppressed according to the NSDUH suppression rules.

Similar to the results for SMI in Tables 5 and 6 in Chapter 4, the results in Tables 14 and 15 indicate that in the smaller 2008-2012 MHSS, AMI estimates among arrestees are somewhat different from those among parolees or probationers (although almost all of these estimates would be suppressed under NSDUH precision-based suppressions rules). However, in the larger 2008-2012 NSDUH, their estimates are similar to those of parolees or probationers. The lack of precision of the estimates in the smaller dataset and the similarity among the estimates in the larger dataset suggests that the differences in estimates between arrestees and parolees or probationers may be due to the small sample sizes of those subpopulations represented in the 2008-2012 MHSS. Therefore, no further analyses of the parolee or probationer subpopulation were conducted for AMI estimation.

Among all the subpopulations represented in Table 15, AMI direct estimates increase as past month K6 score increases and in the presence of lifetime depression and lifetime anxiety, providing further evidence that these mental health indicators would be good predictors of AMI in regression models.

Table 15. Gold-Standard AMI, by Mental Health Indicator and Criminal Justice Status Among Adults 18 or Older, 2008-2012 MHSS

	Criminal Justice Status														
	All Adults			On Probation or Parole†			Arrestee, Not on Probation or Parole			Arrestee, on Probation or Parole‡			Not Arrestee, Not on Probation or Parole		
Mental Health Indicator	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**	**Estimate**		**SE**
Total	18.0	%	0.93	*22.5*	%	*5.40*	*36.8*	%	*8.16*	28.2	%	4.63	17.4	**	0.91
Past Month K6 Score															
< 8	11.9	%	1.06	*11.7*	%	*3.90*	*27.3*	%	*10.00*	*19.0*	%	*4.96*	11.6	%	1.03
8-12	45.7	%	2.76	*21.6*	%	*12.93*	*56.4*	%	*17.08*	*34.5*	%	*13.82*	46.2	%	2.67
> 12	65.9	%	3.34	*51.1*	%	*10.52*	*82.9*	*	*8.23*	*56.9*	%	*8.88*	67.1	%	3.69
Lifetime Depression															
No	12.9	%	1.00	12.6	%	3.60	*37.8*	*	*9.38*	23.0	%	4.69	12.4	**	1.01
Yes	52.3	%	2.59	*74.6*	%	*7.07*	*30.3*	*	*13.94*	*59.2*	%	*10.16*	51.7	%	2.53
Lifetime Anxiety															
No	14.7	%	0.93	14.7	%	3.97	*33.8*	%	*8.80*	22.6	%	4.63	14.2	**	0.92
Yes	53.2	%	3.31	*63.4*	%	*12.64*	*56.7*	%	*20.42*	*61.0*	%	*10.92*	52.5	%	3.30
Past Month K6 Score, Lifetime Depression, and Lifetime Anxiety															
< 8, No Lifetime Depression or Lifetime Anxiety	9.1	%	1.13	*9.5*	%	*3.61*	*29.7*	%	*10.76*	*18.9*	%	*5.48*	8.7	%	1.13
> 12, with Lifetime Depression and Lifetime Anxiety	86.4	%	3.33	*87.0*	%	*9.33*	*65.2*	%	*24.77*	*85.0*	%	*9.19*	86.7	%	3.71
Global Assessment of Functioning Score															
≤ 50	88.2	%	3.68	*68.1*	%	*12.26*	*83.6*	%	*12.21*	*72.3*	%	*9.21*	*91.0*	%	*4.18*
Sample Size	5,653			195			121			316			5,323		

AMI = any mental illness; K6 = Kessler 6-item distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SE = standard error.

*Difference between this estimate and estimate from comparison group † is statistically significant at the 0.05 level.

**Difference between this estimate and estimate from comparison group ‡ is statistically significant at the 0.05 level.

Note: Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules.

Dataset = 2008-2012 MHSS clinical data.
Analysis weight = MHFNLWGT.

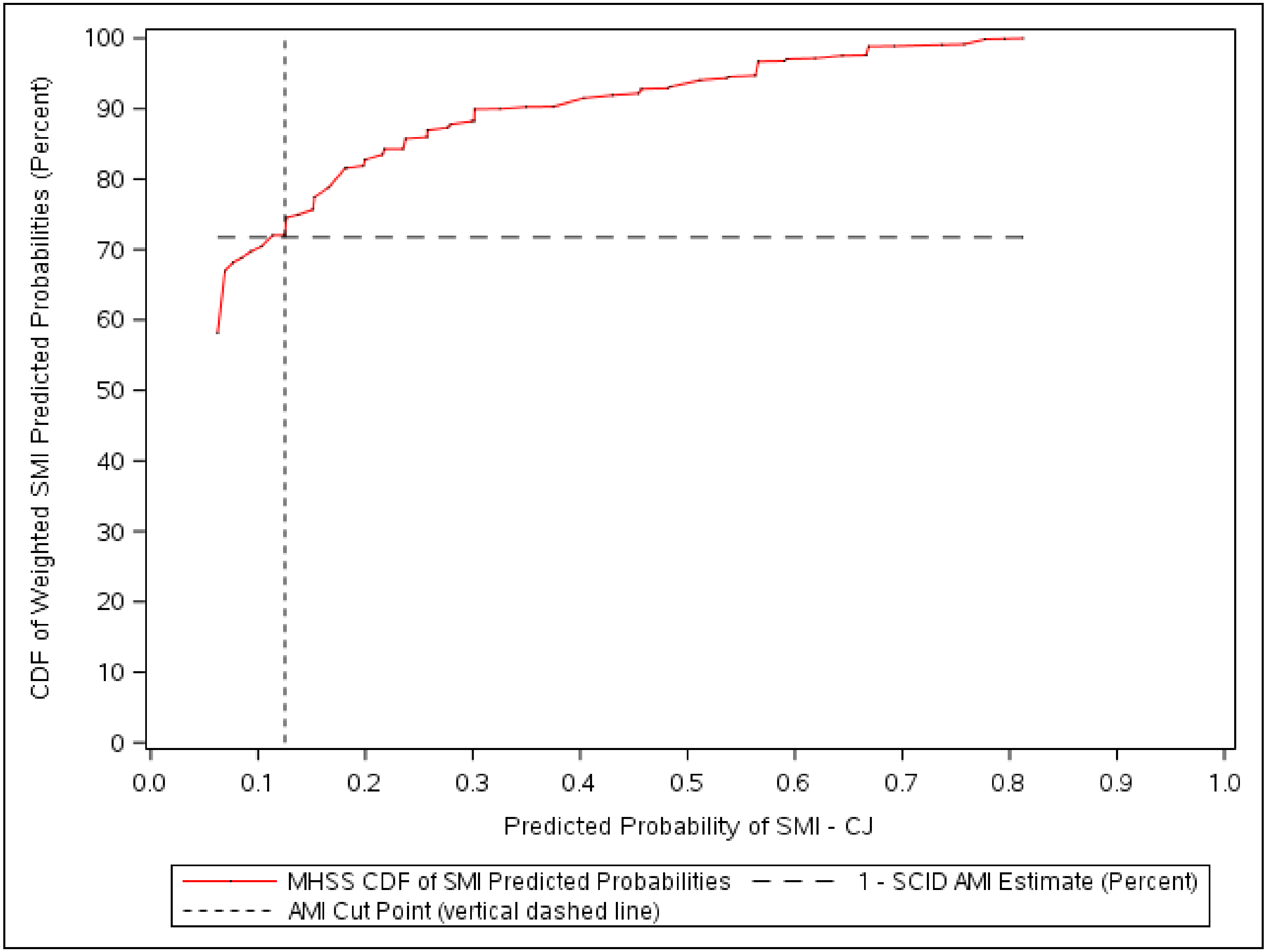
Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

5.2 ROC Analysis of AMI Models

Using the first method for AMI estimation described in Section 5.1, a separate cut point for AMI was determined for the predicted probabilities based on the SMI model K2DA (see Table 7 in Section 4.3 for details of this model). Predicted probabilities (nominally of SMI) were obtained by applying this model to the CJ subsample of the 2008-2012 MHSS data. Then, an AMI cut point was determined so that the predicted probabilities from the model could be dichotomized into a binary cut point AMI estimator that was (approximately) unbiased (i.e., the AMI cut point estimate was as close as possible to the gold-standard AMI direct estimate). Figure 3 illustrates this, which shows the CDF of SMI predicted probabilities based on SMI model K2DA, and demonstrates how the AMI cut point was determined at the point in the CDF equal to 1 minus the SCID direct estimate of AMI expressed as a percentage.

Obviously, the AMI cut point would be lower than the SMI cut point, so more predicted probabilities would meet or exceed the AMI cut point (i.e., those respondents would be predicted as having AMI) in keeping with the less restrictive definition of AMI compared to SMI.

Figure 3. CDF of Weighted SMI Predicted Probabilities Based on SMI Model K2DA for Criminal Justice Subpopulation, 2008-2012



AMI = any mental illness; CDF = cumulative distribution function; MHSS = Mental Health Surveillance Study; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

Using the second method for AMI estimation, separate prediction models for AMI were developed. Weighted logistic regression models were fitted in the CJ subsample of the 2008-2012 MHSS (n = 316). In these models, the response variable was the binary gold-standard AMI variable **SCID_AMI**, and predictor variables included combinations of the four NSDUH variables described in Section 3.4 (i.e., K6SCMON [past month K6 score], K6SCMON2 [alternative past month K6 score], DEPRSLIF_U [lifetime depression], and ANXDLIF_U [lifetime anxiety]). Previous experience using the MHSS indicated that age should be included in the model to reduce the bias of estimates within age groups (CBHSQ, 2014a). Thus, different recoded versions of the age variable were included in models. Many models with combinations of the predictor variables were run, but two models had substantially lower TCE values in the ROC analysis. These two models included only a single predictor variable: either K6SCMON or K6SCMON2. Given the disadvantages of a single-variable model (discussed below), and to assess the effect of the other predictor variables, ROC analyses were conducted on models that included K6SCMON2 and different combinations of DEPRSLIF_U, ANXDLIF_U, and AGE30 (Table 16).

Table 16 displays the ROC analysis of SMI-specific model K2DA (with different cut point for AMI) and five AMI-specific models. Four of the AMI models include K6SCMON2 and all combinations of DEPRSLIF_U and ANXDLIF_U, and the fifth model includes K6SCMON2, DEPRSLIF_U, ANXDLIF_U, and AGE30. The response variable of SMI model K2DA is the binary gold-standard SMI variable **SCID_SMI**, and the SMI cut point for this model is 0.257765 (see Table 8 in Section 4.4). As expected, the AMI cut point for this same model is much lower (i.e., 0.125046). The AMI cut points for the remaining AMI-specific models are much higher (i.e., 0.258187 to 0.332320) where the response variable is the binary gold-standard AMI variable **SCID_AMI**.

Table 16 shows that the AMI model including only K6SCMON2 has the smallest TCE rate (i.e., 25.75). However, in spite of the slightly smaller TCE value of this model, it may be too parsimonious to discriminate accurately between the presence and absence of AMI. In particular, a model that includes only a single predictor variable with 18 distinct levels results in a CDF of predicted probabilities with at most 18 distinct levels that may be too coarse to discriminate between the presence and absence of AMI. Therefore, this AMI model was not considered further. Note that the TCE rates among the AMI models in Table 16 are roughly double those among the SMI models in Table 7. This is because the prevalence estimates of AMI are roughly double those of SMI (i.e., larger prevalence estimates allow for higher false positive and negative rates).

Table 16 also demonstrates that the TCE rate of SMI model K2DA (i.e., 28.10) was slightly smaller than those of all remaining models, except the model including K6SCMON2 and DEPRSLIF_U (i.e., TCE = 27.38). Considering that the difference in TCE rates between the SMI model K2DA and the other AMI-specific models is not large, and consistency problems between SMI and AMI estimates can occur if a separate AMI model is used, the SMI model K2DA (with a different cut point) was used to provide the AMI cut point estimator. By using this approach to estimate AMI (using the SMI model and determining a different cut point for AMI), it is not possible to determine hybrid cut points for the AMI estimator, since there are no AMI-specific predicted probabilities that can be used to obtain the mean predicted probability.

Table 16. ROC Statistics of AMI Cut Point Estimates Based on Selected Candidate Models for Criminal Justice Population

Model	Sample Size	Pop Size (1,000s)	Cut Point	Gold-Standard Estimate	Cut Point Estimate	SE	False Positive Rate	False Negative Rate	Bias	TCE Rate	Sensitivity	Specificity	AUC
K2DA (SMI)	316	8,699	0.125046	28.24	27.97	3.358	13.92	14.18	-0.2654	28.10	0.498	0.806	0.652
K2	316	8,699	0.271983	28.24	29.13	2.721	13.32	12.43	0.8929	25.75	0.560	0.814	0.687
K2D	316	8,699	0.258187	28.24	27.59	3.395	13.37	14.01	-0.6469	27.38	0.504	0.814	0.659
K2A	316	8,699	0.332320	28.24	28.09	3.282	14.05	14.21	-0.1524	28.26	0.497	0.804	0.651
K2DA	316	8,699	0.309195	28.24	28.09	3.282	14.05	14.21	-0.1524	28.26	0.497	0.804	0.651
K2A30	316	8,699	0.306084	28.24	27.97	3.293	14.00	14.27	-0.2662	28.27	0.495	0.805	0.650

AMI = any mental illness; AUC = area under curve (i.e., average of sensitivity and specificity); CJ = criminal justice; MHSS = Mental Health Surveillance Survey; Pop = population; ROC = receiver operating characteristic; SE = standard error of cut point estimate; SMI = serious mental illness; TCE = total classification error.

Note: Bias = false positive rate – false negative rate; TCE rate = false positive rate + false negative rate.

Response variable = SCID_SMI (gold-standard SMI) for model K2DA (SMI), and SCID_AMI (gold-standard AMI) for all other models.

Predictor variables included in the following models:

K2DA (SMI):	K6SCMON2 + DEPRSLIF_U + ANXDLIF_U
K2:	K6SCMON2
K2D:	K6SCMON2 + DEPRSLIF_U
K2A:	K6SCMON2 + ANXDLIF_U
K2DA:	K6SCMON2 + DEPRSLIF_U + ANXDLIF_U
K2A30:	K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE30

Dataset = CJ subsample of 2008-2012 MHSS clinical data.
Analysis weight = MHFNLWGT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

5.3 Initial Application and Assessment of AMI Estimates in 2016 SPI

In this final step of predicting AMI for the 2016 SPI prisoner data, the SMI prediction model K2DA, developed within the CJ subsample of the MHSS, was selected for extrapolation to the 2016 SPI. In addition, the separate cut point for AMI determined from the CJ subsample of the MHSS (i.e., 0.125046) also was used on the 2016 SPI data. This allowed an AMI cut point estimate to be obtained from the 2016 SPI data. Note that the AMI cut point estimate is independent of the SMI cut point estimate because the AMI cut point is not affected by the final SMI cut point (i.e., they were determined separately).

In addition, AMI cut point estimates were obtained for the same domains within the 2016 SPI data that were used for SMI cut point estimates (see Section 4.7):

- sex (male, female)
- age (18 to 24, 25 to 34, 35 or older)
- race/Hispanic origin (non-Hispanic white, non-Hispanic black, Hispanic, non-Hispanic other)
- marital status (married, widowed, divorced or separated, never married)
- military service ("yes," "no")
- jurisdiction (federal, state)
- offense (violent, property, drug, public order, other).

For comparative purposes, AMI estimates from both the MHSS and NSDUH are presented in Table 17 for only the domains associated with the 2016 SPI.

In Table 17, the first two columns of AMI estimates are associated with the 2008-2012 MHSS. The first column is associated with SCID direct estimates and the second with cut point estimates within the entire CJ population. These two sets of estimates agree pretty closely overall and across most domains (28.2% and 27.9%, respectively). However, all of the domain estimates are highly imprecise, thus the differences in subdomain estimates across the estimation methods and data sources cannot be compared. Almost all of the domain estimates would have been suppressed by the NSDUH suppression rules.

The next column of AMI estimates in Table 17 is associated with the AMI cut point from the MHSS applied to the CJ population of the 2008-2012 NSDUH. The overall estimate agrees with those from the CJ population of the 2008-2012 MHSS (27.7%).

The final column of AMI estimates in Table 17 is associated with the 2016 SPI. These estimates are based on SMI model K2DA and the AMI cut point from the MHSS applied to the 2016 SPI data. The estimates are substantially larger than those associated with the MHSS or NSDUH (overall estimate is 35.8%).

Table 17. AMI Estimates Based on SMI Model K2DA for Different Data and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI

Dataset	**2008-2012 MHSS**		**2008-2012 NSDUH**	**2016 SPI**
Subpopulation	**Criminal Justice**		**Criminal Justice**	**All Prisoners**
Estimator	**SCID**	**MHSS Cut Point**	**MHSS Cut Point**	**MHSS Cut Point**
Cut Point		**0.12505**	**0.12505**	**0.12505**
Sample Size	**316**	**316**	**15,400**	**24,848**
Demographic Variable				
Total	28.2%	27.9%	27.7%	35.8%
Sex				
Male	26.8%	22.9%	21.4%	33.9%
Female	*33.0%*	*44.9%*	43.7%	61.7%
Age				
18-24	*28.4%*	*18.3%*	22.4%	33.5%
25-34	*23.8%*	*42.7%*	27.7%	37.7%
35 or Older	*30.5%*	*25.5%*	31.7%	35.2%
Race/Hispanic Origin				
Non-Hispanic White	*32.0%*	*31.3%*	34.1%	46.0%
Non-Hispanic Black	*30.3%*	*31.7%*	18.9%	26.9%
Hispanic	*11.9%*	*22.7%*	18.8%	29.7%
Non-Hispanic Other	*29.8%*	*8.4%*	25.9%	45.5%
Marital Status				
Married	*26.2%*	*22.6%*	27.0%	34.7%
Widowed	*23.0%*	*1.9%*	*41.3%*	34.6%
Divorced or Separated	*38.0%*	*36.0%*	35.7%	39.6%
Never Married	27.5%	31.1%	24.7%	34.7%
Military Service				
Yes	*31.4%*	*37.9%*	28.0%	40.2%
No	28.1%	27.4%	27.6%	35.5%

(continued)

Table 17. AMI Estimates Based on SMI Model K2DA for Different Data and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI (continued)

Dataset	**2008-2012 MHSS**		**2008-2012 NSDUH**	**2016 SPI**
Subpopulation	**Criminal Justice**		**Criminal Justice**	**All Prisoners**
Estimator	**SCID**	**MHSS Cut Point**	**MHSS Cut Point**	**MHSS Cut Point**
Cut Point		**0.12505**	**0.12505**	**0.12505**
Sample Size	**316**	**316**	**15,400**	**24,848**
Demographic Variable				
Jurisdiction				
Federal	~	~	~	21.0%
State	~	~	~	37.9%
Offense				
Violent	~	~	~	37.0%
Property	~	~	~	41.3%
Drug	~	~	~	31.9%
Public Order	~	~	~	32.3%
Other	~	~	~	:

AMI = any mental illness; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness; SPI = Survey of Prison Inmates.

:Not calculated.

~Not applicable.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. Standard errors of estimates in this table are contained in Table B.17 in Appendix B.

Datasets = 2008-2012 MHSS clinical data for MHSS estimates, 2008-2012 adult NSDUH data for NSDUH estimates, and 2016 SPI data for SPI estimates.
Analysis weights = MHFNLWGT for MHSS estimates, ANALWT for NSDUH estimates, and WT_FINAL for SPI estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012; and Bureau of Justice Statistics, Survey of Prison Inmates, 2016.

The SPI estimates in Table 17 show that females have higher rates of AMI than males, and non-Hispanic whites have higher rates than non-Hispanic blacks and Hispanics. These results are consistent with those among the CJ population in the NSDUH, and the pattern of these results (i.e., higher versus lower) are consistent with SMI rates in Table 13. The SPI estimates in Table 17 show similar rates of AMI among the three age groups, and the SPI estimates show lower rates among federal prisoners. Both findings are consistent with the pattern of SMI rates in Table 13.

The three overall AMI estimates from the MHSS or NSDUH data are consistently around 28%, lower than the SPI estimate of 36%. The potential reasons for these differences and a brief discussion on the limitations of AMI estimation in the 2016 SPI follow in the next section.

5.4 AMI Estimation in 2016 SPI: Considerations

There are several reasons why the overall estimates of AMI range from 27.7% to 35.8% across all four sets of estimates in Table 17. Because the AMI cut point estimator is based on SMI model K2DA but with a different cut point, most of the issues about SMI estimation in the 2016 SPI discussed in Section 4.7 apply similarly to AMI estimation. Also, AMI estimation has additional issues associated with it that do not apply to SMI estimation. However, the issues that are similar to those associated with SMI estimation are discussed first below.

First, the *sample size* of the CJ subsample of the 2008-2012 MHSS is small (n = 316). However, this is the data from which the K2DA model was developed and the MHSS cut point for AMI determined. Figure 3 in Section 5.2 shows how sample size can affect the smoothness of the CDF of predicted probabilities associated with this model. CDFs that are not smooth due to small sample sizes make it difficult to precisely determine an appropriate cut point. Figure 3 illustrates this with respect to the AMI cut point. In spite of this, all MHSS and NSDUH estimates of AMI agree well. In addition, the predicted probabilities and estimates of the model parameters will be less precise due to the small sample size.

Second, *extrapolation* of a model (and cut point) from a subsample of a larger sample to the larger sample itself appears to be reasonable if the size of the subsample is sufficiently large (e.g., Figure 1 suggests this was true for the 2008-2012 MHSS). However, these SPI estimates are based on extrapolation of a model (and cut point) from a small subsample of a clinical sample from one survey (i.e., CJ subsample of 2008-2012 MHSS) to an entirely different survey (i.e., the 2016 SPI). The two surveys possess different methodological characteristics (e.g., sampling frame, data collection methods). In particular, the sampling frames of the 2008-2012 NSDUH (from which the 2008-2012 MHSS was subsampled) and the 2016 SPI have been designed to have almost no overlap. The 2008-2012 NSDUH specifically excluded prisoners from the sampling frame, whereas the 2016 SPI included only prisoners in its sampling frame. (However, it is possible that someone in the CJ population in 2008 to 2012 could have been in prison in 2016). In addition, the surveys collect data in different settings (i.e., NSDUH typically in respondents’ homes, SPI in prisons) using different modes (i.e., ACASI for NSDUH and CAPI for SPI), which likely result in different reporting behavior. These survey design differences could account for some of the reason why the SPI AMI estimate of 35.8% is substantially larger than those associated with the MHSS or NSDUH (i.e., 27.7% to 28.2%).

Third, there is *limited information* in the 2016 SPI available for the development of SMI prediction models (used to obtain AMI estimates with a different cut point). The 2016 SPI did not collect information on the past year K6 scale, the WHODAS impairment measure, past year suicidality, or past year MDE. All of these variables were used to construct predictor variables for the 2012 NSDUH prediction model in the MHSS (which also was used to obtain AMI estimates with a different cut point). The 2016 SPI collected information on the past month K6 scale, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety, all of which were used in the SMI model K2DA. However, these SPI mental health indicators might not predict AMI as accurately as the variables available in NSDUH.

Fourth, the *populations* defined by criminal justice status (even if subsetted to parolees or probationers only) in the 2008-2012 NSDUH and actual prisoners in the 2016 SPI likely differ at least to some extent. Some differences may be due simply to the different time frames of the two data sources: 2008 to 2012 for NSDUH and 2016 for SPI. Most significantly, the NSDUH population is currently not incarcerated, while the SPI population is. Of course, parolees and possibly some probationers and arrestees also were formerly incarcerated, so that difference likely would not be as large as, for example, between prisoners and the non-CJ population. The surveys actually cover many of the same populations but interview them at different times (and living situations and modes) during their lives. However, if the two populations differ, it does not automatically mean that a prediction model fit to one population will not apply to a different population. For example, SMI model K2DA is based on past month K6 scores, lifetime depression, and lifetime anxiety. These terms representing psychological distress, lifetime depression, and lifetime anxiety may apply similarly to both populations. As another example, the 2012 NSDUH prediction model developed in the 2008A-2012 MHSS is assumed to work equally well for the CJ and non-CJ populations. Based on the 2012 NSDUH prediction model, the estimates of AMI for the CJ population are higher than those from the non-CJ population simply because respondents in the CJ subsample tended to have higher levels of psychological distress, impairment, MDE, or thoughts of suicide. Therefore, the higher rates of AMI estimated in the 2016 SPI may be a consequence of prisoners experiencing higher levels of psychological distress, depression, or anxiety. However, there is no way to test this assumption with the current information available, but it does appear to be consistent with the findings of Bronson and Berzofsky (2017). Rates of psychological distress, depression, or anxiety are higher in the prisoner population than among the general population.

Finally, the AMI estimator is based on a prediction model for SMI but with a different cut point. However, the SMI prediction model is developed to predict SMI specifically and may not be as effective at identifying cases that have other levels of mental illness (e.g., AMI). Also, because the mean of the predicted probabilities of the SMI prediction model will result in an estimate of SMI (i.e., not AMI), hybrid cut points for AMI cannot be determined. Therefore, the AMI estimator for the 2016 SPI also depends on the AMI cut point determined in the CJ subsample of the 2008-2012 MHSS.

5.5 Selection of Final AMI Model in the 2016 SPI

Analyses in this chapter have shown that psychological distress as measured by the past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety are associated with AMI among the CJ population of the 2008-2012 MHSS and NSDUH. Analyses have also shown that the SMI model K2DA (which includes terms representing alternative past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety) provides cut point estimates (using a different cut point for AMI) within the CJ subsample of the 2008-2012 MHSS that are similar to gold-standard SCID direct estimates of AMI overall and across most domains. The same model and AMI cut point extrapolated to the CJ subsample of the adult 2008-2012 NSDUH provides similar AMI estimates (overall and across most domains). Therefore, unlike in the case of SMI estimation, a separate hybrid cut point does not necessarily need to be determined for AMI. Of course, given the AMI estimation approach followed, it would not have been possible to do so anyway.

If we can assume that the association between past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety are associated with AMI among the prisoner population similarly to the observed association among the CJ population of the 2008-2012 MHSS and NSDUH, then SMI model K2DA with a separate cut point for AMI is likely to provide reasonable predicted probabilities of AMI among the prisoner population. As noted in Sections 4.8 and 5.4, the higher predicted probabilities observed among the prisoner population are likely due to the fact that this population differs somewhat from the CJ population not currently incarcerated. Therefore, AMI estimates can be obtained for the 2016 SPI from the SMI prediction model expressed as—

$$\begin{aligned}\text{logit}(\hat{\pi}) \equiv \log\left[\hat{\pi}/(1-\hat{\pi})\right] = &-2.7108537 + 0.1093374 * X_{km} + 1.5438971 * X_{dep} \\ &+0.7763546 * X_{anx},\end{aligned}$$

where $\hat{\pi}$ is the estimated probability that an adult had SMI, and the predictor variables X_{km}, X_{dep}, and X_{anx} are defined as—

X_{km} = *Alternative Past Month K6 Score:* past month K6 score of less than 8 recoded as 0; past year K6 score of 8 to 24 recoded as 1 to 17 [i.e., **K6SCMON2**].

X_{dep} = *Lifetime Depression*: coded as 1 if “yes”; coded as 0 otherwise [i.e., **DEPLIF**].

X_{anx} = *Lifetime anxiety*: coded as 1 if “yes”; coded as 0 otherwise [i.e., **ANXLIF**].

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

$$\text{SMIPP} = 1 \Bigg/ \left(1 + \exp\left[-\begin{pmatrix}-2.7108537 + 0.1093374 * X_{km} + 1.5438971 * X_{dep} \\ +0.7763546 * X_{anx}\end{pmatrix}\right]\right).$$

If SMIPP is greater than or equal to 0.125045755 (the AMI cut point from the CJ subsample of the 2008-2012 MHSS data), then the respondent is predicted as having past year AMI; otherwise, the respondent is predicted as not having past year AMI.

6. Limitations of the Model-Based Estimates

6.1 Alternative Approaches for Estimating SMI and AMI Among Prisoners

The most rigorous and accurate approach to estimating SMI and AMI among prisoners would be to administer clinical diagnostic interviews to a large, representative sample of prisoners (e.g., SPI). However, constraints related to costs, level of effort, and burden on correctional facilities have made this approach infeasible for samples as large as the 2016 SPI (n = 24,848).

An alternative approach is to include validated scales assumed to measure SMI and AMI reasonably well in the SPI questionnaire. However, although validated scales for the general population can measure the individual disorders that are used to determine SMI and AMI, none have been validated on a prisoner population, and the amount of interview time required to administer enough of these scales to cover "any disorder" would be beyond the survey's capability. BJS needs to address several other high-priority topics in its prisoner surveys. Additionally, during 2009 and 2010, BJS collaborated on a project with the National Institute of Mental Health (NIMH) and Dr. Robert Trestman to develop and validate a scale to include in BJS's national prisoner and jail inmate surveys to produce direct, national-level estimates of SMI and specific psychological disorders prevalent in correctional systems and of clinical and operational significance in correctional settings. The project involved a sample of about 310 inmates selected from three state prison facilities and three jail facilities in one state. A composite 62-item questionnaire that included diagnostic questions for major depression, bipolar depression (Type I and Type II), generalized anxiety disorder, panic disorder, serious phobias, post-traumatic stress disorder, and borderline personality disorder (Axis II) was fielded among the sample. The 62 questions were chosen from instruments that were previously proven to be valid and reliable[20] and were selected based on feedback from experts and stakeholders of various disciplines. The study sought to reduce the total number of questions to the smallest number that best predict SMI and the specific psychological disorders that best represent SMI. This core set of questions would represent the final scale to include in BJS's inmate surveys. A clinical interview using the Structured Clinical Interview for DSM-IV-TR-SCID I and Structured Clinical Interview II for DSM-IV (borderline personality disorder section only) was conducted to assess current and past Axis I psychiatric disorders among the sample of 310 and was used as the gold standard. A 16-item scale was developed by using statistical techniques to determine which of the 62 items best predicted disorders as measured by the gold-standard clinical interview. However, the final recommendation was that the 16-item scale should undergo further testing to demonstrate its reliability and validity among a sample of prisoners and jail inmates in other states (Trestman & Shelton, 2010). Given this recommendation, BJS decided not to include the 16-item scale in its inmate surveys, including the 2016 SPI, until the results could be replicated. A follow-up study has not yet been conducted at this point. It should be noted that this method for estimating SMI uses a different definition of SMI than SAMHSA's definition. SAMHSA defines SMI as *any* mental disorder that causes serious impairment. The methods described for

[20] Instruments included the Composite International Diagnostic Interview, National Inmate Survey, Structured Clinical Interview for DSM-IV Screening Module, Primary Cary Post-Traumatic Stress Disorder Screener, Iowa Personality Disorder Screen, and K6.

the BJS-NIMH-Trestman study selected certain mental disorders as SMI without the use of a specific impairment criterion.

SAMHSA's model-based approach should also be considered in future BJS prisoner surveys. A subsample of main survey (SPI) respondents would be selected for clinical interviews, and a set of short scales would be included in the main survey interview. Therefore, this subsample would have both gold-standard clinical assessments *and* indicators of mental illness (e.g., K6 scale) collected from the main interview. This would allow a predictive model to be developed matching the two sources of information. The model would be applied to all survey respondents to obtain nationally representative estimates of SMI and AMI among prisoners. However, conducting clinical interviews in a prison setting, along with the main SPI survey interviews, would place an additional substantial burden on the sampled facilities and their staff that is not feasible currently. Also, BJS has concerns about the impact this could have on response to the main SPI study, which is fielded periodically given the scale and costs.

The first and third approaches described above for estimating the prevalence of mental illness among prisoners have not been attempted. The second approach has not yet resulted in a scale that has been fully tested to be reliable and valid, and no gold-standard clinical assessments exist for prisoner respondents in the 2016 SPI. Nevertheless, given the critical need for valid estimates of mental illness among prisoners, BJS is exploring the possibility of developing estimates from existing data, taking advantage of the nationally representative sample of clinical assessments available from the 2008-2012 MHSS and the 2016 SPI. A SMI prediction model was developed within a proxy subpopulation for prisoners, and the CJ subpopulation represented by the 2008-2012 MHSS. The CJ subpopulation consists of parolees, probationers, and arrestees, and is assumed to have characteristics similar to those of prisoners. The SMI prediction model was extrapolated to the 2016 SPI data to produce estimates of SMI and AMI in the 2016 SPI.

The following sections present assessments of the methodology, in general and specifically focusing on the NSDUH CJ model and the 2016 SPI estimates.

6.2 General Limitations of Model-based Estimates of SMI and AMI

Using the general model-based method to estimate SMI and AMI presents several potential limitations, learned from experiences based on the 2008-2012 MHSS (CBHSQ, 2014a). These limitations may arise due to potential impacts from the following factors: weights, sample size, covariate selection, model parsimony, and estimation of different levels of mental illness.

First, the model-based method assumes that the clinical measures of SMI and AMI on the MHSS are close to the "true" measures of SMI and AMI and can be considered the gold standards for the purposes of this effort. Therefore, model-based measures of bias and error use the gold-standard measures as the reference measures. If any differences exist between gold-standard and "true" measures, model-based assessments do not account for these differences.

Because a primary goal is to provide nationally representative estimates of SMI and AMI, model-based analyses are based on *weighted* data. If the sample design used to select the subsample of respondents to undergo clinical interviews is optimized in any way (e.g., by the

Neyman optimal allocation method; Cochran, 1986), then this could increase the variability in the subsample weights. Some weights might be large relative to others, and large weights in the clinical subsample have the potential to present problems. For example, a large weight "straddling" the neighborhood of a potential cut point could make it difficult to come close to equalizing false positive and false negative weighted counts. If the cut point is placed on one side of the weighted predicted probability, it could result in a fairly large bias in one direction, but if placed on the other side, it could result in an almost equally large bias in the other direction.

If the *sample size* of the clinical subsample is small, this presents limitations associated with (1) direct gold-standard estimates, (2) model-based analyses, and (3) subsequent model-based estimates.

At the respondent level, the gold-standard assessment may be assumed to be a true assessment of mental illness, but at the (nationally representative) aggregated level, prevalence estimates may be subject to large design-based sampling error if the sample size of the clinical subsample is small (particularly at the domain level, e.g., widowed respondents). If the design-based sampling error of the gold-standard estimates is large, then this will directly affect subsequent model-based analyses because the model is built around a response variable made up of gold-standard responses (i.e., the gold-standard responses act as the "benchmark" reference for the model).

Smaller sample sizes also will result in larger model-based error (i.e., the standard errors of the beta estimates of the model will be larger). Perhaps more importantly though, if the sample size of the clinical subsample is small, then model specification error may occur (i.e., it may be difficult to identify the "best" prediction model from the available data). For example, in the 2008A (i.e., Sample A) MHSS (n = 759), the available data only allowed a model with 2 degrees of freedom (DF) to be developed. The terms in this model included alternative versions of the K6 and WHODAS scores. In comparison, the much larger 2008A-2012 MHSS (n = 4,912) allowed a 5 DF model to be developed. This model included the two terms in the 2008A model, terms for past year thoughts of suicide and past year MDE (both of which resulted in a fairly large reduction in the TCE rate), and an age variable that substantially reduced bias within some age groups (CBHSQ, 2014a).

Small sample sizes may also result in a CDF of predicted probabilities that is not smooth. This could be exacerbated if there is a fair amount of variability in the weights, which could result in cut point error (i.e., it might be difficult to determine a precise cut point). (See Figure 2 in Section 4.1.5 for an example of this.)

All four sources of error, specifically design-based sampling error in gold-standard estimates, model-based error, model specification error, and cut point error, due to small sample sizes will contribute to error in the model-based prevalence estimates.

Potential limitations may arise from the *covariates selected* for the model. Outcome variables closely related to mental illness that also form domains of interest (e.g., receipt of mental health services during the past year) should be avoided. If such a variable is included in the model, then SMI prevalence would tend to be overpredicted within some domain(s) formed from that variable (CBHSQ, 2014a). For example, because the 2012 NSDUH prediction model included past year thoughts of suicide and past year MDE (because their inclusion resulted in a fairly large reduction in the TCE rate), analyses of SMI within the domains “had suicidal thoughts during the past year” and “had past year MDE” were no longer feasible due to the overprediction problem.

Another set of variables associated with sociodemographic domain variables should be avoided. Inclusion of such variables would cause the relationship between SMI and that domain to be fixed (which might not be desirable if the model were used to estimate SMI in later years). A domain variable might be included in the model as a predictor variable only if it brought with it substantial benefits. For example, in the 2012 NSDUH prediction model, an age variable was included because it substantially reduced bias within some age groups.

Model parsimony is usually a desirable property if a model will be extrapolated to a larger dataset. For example, a model that is too enriched may be tailored too closely to the particular properties of the subsample in which it is developed, and hence may not extrapolate well. Considering the limitations associated with the types of predictor variables to include in a model, final prediction models do tend to be parsimonious (e.g., the 2012 NSDUH prediction model has only 5 DF). However, if a model is too parsimonious then this might result in cut point error. For example, consider a model with only one predictor variable, the alternative K6 score. This predictor variable has only 18 distinct values, and hence the model will output at most 18 distinct predicted probabilities. Therefore, the CDF of predicted probabilities will be coarsely gradated. Because cut points can theoretically occupy any value within the interval between gradations (i.e., where the interval is open at the bottom end and closed at the top end), this could result in substantial cut point error.

Limitations may also occur if there is interest in estimating *levels of mental illness* other than SMI (e.g., AMI). One approach is to develop polytomous regression models (i.e., where the response variable has more than two levels). This approach is usually not considered because it is complicated and optimizes over all levels of mental illness represented in the polytomous response variable, whereas typically SMI is of primary interest, and other levels (e.g., AMI) are of secondary interest. In a more simple approach, separate models for each mental illness variable can be developed. However, consistency of estimates resulting from the different models is not guaranteed. The preferred approach is to use a single SMI prediction model with different cut points to obtain estimates of SMI, and even AMI. This approach is simple, provides consistent estimates, and appeared to work well in the 2008-2012 MHSS (CBHSQ, 2014a). However, one potential limitation of this approach is that an SMI prediction model is developed specifically to predict SMI. Therefore, it may not effectively identify cases that have other levels of mental illness. Another limitation is that the mean of the predicted probabilities of the SMI prediction model will result in an estimate of SMI (i.e., not AMI), and in the absence of predicted probabilities specifically for AMI, it will not be possible to determine hybrid cut points for AMI.

However, given all of the limitations in the general model-based method described above, it is important to realize that the purpose of this method is to provide prevalence estimates of SMI and AMI at the *aggregated* level (i.e., where minimization of bias at the aggregated level is important). The model-based method is *not* designed to be used as an individual diagnostic test (i.e., where it is important to have high rates of sensitivity and specificity, as well as positive predicted value and negative predicted value).

6.3 Additional Limitations of the Methodology Used to Estimate SMI and AMI Among Prisoners

In addition to the limitations of the general model-based estimation approach discussed in Section 6.2, there are other limitations that are specific to the methodology developed to estimate SMI and AMI among prisoners.

The first specific limitation of the methodology arises from the *small sample size* of the clinical sample suitable for this analysis. (See Section 6.2 for a discussion of the impact of small sample size on the general model-based approach.) Although the size of the entire 2008-2012 MHSS clinical sample is 5,653, only a small portion of the clinical sample was part of the subpopulation of interest. This subsample consisted of 316 adult respondents who self-reported as parolees, probationers, or arrestees during the past 12 months. These respondents were assumed to possess characteristics similar to those of prisoners. As noted in Section 6.2, a sample size this small may result in substantial levels of design-based sampling error in gold-standard estimates and errors associated with modeling. A second specific limitation arises from the *limited overlapping informatio*n between the NSDUH and the 2016 SPI that could be used in the development of SMI prediction models. The 2016 SPI collected information on past month psychological distress (K6 scale), lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety. However, it did not collect information on past year K6 scale, WHODAS impairment measure, past year suicidality, or past year MDE, which were used to construct predictor variables for the 2012 NSDUH prediction model in the MHSS.

A third specific limitation of the methodology arises from the need to use a *different population* (i.e., CJ subpopulation) as a proxy for the prisoner population. This means that the *populations* represented by criminal justice status (even if subsetted to parolees or probationers only) in the 2008-2012 NSDUH and actual prisoners in the 2016 SPI are likely to differ to some extent. The major different between the two populations is that the first is currently not incarcerated while the second is. Because parolees and possibly some probationers and arrestees also were formerly incarcerated, the differences between the NSDUH CJ population and prisoners are not expected to be as large as between prisoners and the non-CJ population. However, the two data collections have other methodological differences, such as when they were interviewed (during or after incarceration), data collection period, mode of data collection (ACASI for NSDUH and CAPI for SPI), the questions themselves, and questionnaire context. Also, the NSDUH *excludes* current prisoners from the sampling frame, whereas the 2016 SPI included *only* prisoners in its sampling frame. Thus, this effort is predicated on the assumption that the model (and cut point) developed based on the CJ subsample of the 2008-2012 MHSS would hold for a different survey (i.e., 2016 SPI). It is possible that some of these methodological differences may not have a strong impact. For example, the different time frames of the data collections (i.e., NSDUH in 2008-2012 and SPI in 2016) are unlikely to have much of an effect because estimates of SMI and AMI appear to be fairly stable over time. For example, in the NSDUH, the SMI prevalence estimate among adults is 3.9% in 2008-2012 and 4.2% in 2016. The AMI prevalence estimate among adults is 18.1% in 2008-2012 and 18.3% in 2016. Another factor to take into account in any model-based estimation system is that models are valid across time periods, as discussed below.

While it is important to recognize the differences in populations covered, survey methods, and time frames between the NSDUH and SPI, those differences do not necessarily cause large bias in the model-based estimates, even if prevalence rates for certain measures differ greatly. The underlying assumption in extrapolating the NSDUH-based models to the SPI sample is that the *model prediction* is similar across the surveys. In other words, it is accepted (and shown by the data) that for many indicators, the surveys produce different prevalences. Those differences do not represent bias in the SMI and AMI estimates, as long as the *relationships* between SMI and AMI with predictors are similar in both samples. For example, regardless of the mean K6 score in each survey, in both surveys, a higher K6 score is expected to be associated with a higher probability of having SMI or AMI.

In addition, even if the point estimates of SMI and AMI have some error, patterns of estimates among domains such as sex and race/Hispanic origin are consistent with those reported among the CJ population in the NSDUH and in three prior BJS reports (Bronson & Berzofsky, 2017; James & Glaze, 2006; Ditton, 1999). Patterns of estimates for different age groups are consistent with those reported in Bronson and Berzofsky (2017), and patterns of estimates for federal and state prisoners are consistent with those reported in James and Glaze (2006) and Ditton (1999).

6.4 Assessment of the Final Model

An attempt was also made to specifically address the limitation of not having the same predictive variables in the SPI compared to the NSDUH (e.g., past year K6, past year MDE, past year thoughts of suicide as discussed in Section 6.3). The goal was to assess the importance of the terms in the 2012 NSDUH prediction model that were excluded from SMI model K2DA, which includes alternative past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety. This was done by conducting a ROC analysis to the following six prediction models within the CJ subsample of the 2008-2012 MHSS and comparing this to the original ROC analysis of the 2012 NSDUH prediction model (referred to as **M7** for the rest of this section).[21] The six models developed in the CJ subsample of the 2008-2012 MHSS included—

- **M1**: alternative past month K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety (i.e., SMI model K2DA)
- **M2**: alternative past year K6 score, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety
- **M3**: 2012 NSDUH prediction model, with original betas and cut point from the MHSS
- **M4**: 2012 NSDUH prediction model, with original betas from the MHSS, but cut point determined from the CJ subsample

[21] The 2012 NSDUH prediction model includes five predictor variables: alternative past year K6 score, alternative WHODAS score, serious thoughts of suicide during the past year, past year MDE, and AGE1830 (see Section 2.7 for details).

- **M5**: 2012 NSDUH prediction model, with betas and cut point determined from the CJ subsample
- **M6**: 2012 NSDUH prediction model without AGE1830 variable, with betas and cut point determined from the CJ subsample.

Table 18 presents the results of the ROC analysis of the six models. Based on TCE and absolute bias values (lower is better in both cases), the following general conclusions regarding the CJ population can be drawn:

- A comparison of **M1** and **M2** suggests that alternative past year K6 score is a better predictor variable than alternative past month K6 score.
- A comparison of **M3**, **M4**, and **M5** suggests that for the 2012 NSDUH prediction model, the original betas from the MHSS appear to perform better than those fitted to the CJ subsample data (probably a consequence of the small sample size of the CJ subsample). However, the MHSS cut point extrapolated to the CJ subsample results in a substantial amount of negative bias (so the cut point determined in the CJ subsample itself is to be preferred).
- A comparison of **M2** and **M6** suggests that in addition to alternative past year K6 score, predictor variables representing the WHODAS, past year MDE, and past year thoughts of suicide appear to provide only moderately better predictions than predictor variables representing lifetime self-reported diagnosed depression and lifetime self-reported diagnosed anxiety.
- A comparison of **M5** and **M6** suggests that dropping the AGE1830 variable appears to improve the model substantially.
- A comparison of **M7** with all other models is somewhat of an apples-to-oranges comparison due to the much lower SMI prevalence rate among the general population. A lower prevalence rate tends to reduce the TCE rate and sensitivity, and increase specificity. However, the AUC for this model is squarely within the range of those of the other models.

This analysis suggests that for the CJ population at least, the selected SMI model K2DA appears to provide estimates of SMI that are almost as accurate as those based on the 2012 NSDUH prediction model. However, there would have been a marked improvement if alternative past year K6 score could have been used in the model instead of alternative past month K6 score. Note that the AUC statistics of all of the six models developed within the CJ subsample of the MHSS are similar to the AUC of the original 2012 NSDUH prediction model. In all seven cases, the AUC is in the range 0.70-0.79, which is often interpreted as indicating moderate agreement (Landis & Koch, 1977).

Table 18. ROC Statistics of SMI Cut Point Estimates Based on Selected Candidate Models for Criminal Justice Population, 2008-2012 MHSS

Model	Sample Size	Pop Size (1,000s)	Cut Point	Gold-Standard Estimate	Cut Point Estimate	SE	False Positive Rate	False Negative Rate	Bias	TCE Rate	Sensitivity	Specificity	AUC
M1	316	8,699	0.257765	13.89	14.02	3.037	7.06	6.93	0.1286	13.99	0.501	0.918	0.710
M2	316	8,699	0.305448	13.89	13.48	2.656	6.05	6.46	-0.4117	12.51	0.535	0.930	0.732
M3	255	8,329	0.260574	13.40	11.38	2.852	5.02	7.04	-2.0243	12.05	0.475	0.942	0.708
M4	255	8,329	0.156920	13.40	13.10	2.821	5.94	6.24	-0.3039	12.18	0.534	0.931	0.733
M5	255	8,329	0.315766	13.40	13.22	3.168	6.57	6.75	-0.1797	13.32	0.496	0.924	0.710
M6	255	8,329	0.281476	13.40	14.25	2.940	6.12	5.28	0.8458	11.40	0.606	0.929	0.768
M7	4,912	231,890	0.260574	3.93	3.92	0.270	1.92	1.93	–0.0128	3.84	0.509	0.980	0.745

AUC = area under curve (i.e., average of sensitivity and specificity); CJ = criminal justice; K6 = Kessler 6-item distress scale; MHSS = Mental Health Surveillance Study; NSDUH = National Survey of Drug Use and Health; Pop = population; ROC = receiver operating characteristic; SE = standard error of cut point estimate; SMI = serious mental illness; TCE = total classification error.

Note: Bias = false positive rate – false negative rate; TCE rate = false positive rate + false negative rate.

Response variable = SCID_SMI (gold-standard SMI).

Predictor variables included in the following models:

- M1: alternative past month K6 score, lifetime self-reported diagnosed depression, lifetime self-reported diagnosed anxiety
- M2: alternative past year K6 score, lifetime self-reported diagnosed depression, lifetime self-reported diagnosed anxiety
- M3: 2012 NSDUH prediction model, with original betas and cut point from the MHSS
- M4: 2012 NSDUH prediction model, with original betas from the MHSS, but cut point determined from the CJ subsample
- M5: 2012 NSDUH prediction model, with betas and cut point determined from the CJ subsample
- M6: 2012 NSDUH prediction model without AGE1830 variable, with betas and cut point determined from the CJ subsample
- M7: 2012 NSDUH prediction model in 2008A-2012 MHSS

Datasets = for M1-M2, CJ subsample of 2008-2012 MHSS clinical data; for M3-M6, CJ subsample of 2008A-2012 MHSS clinical data; for M7, 2008A-2012 MHSS clinical data.
Analysis weights = for M1-M2, MHFNLWGT; for M3-M7, MHFAAWGT.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012.

7. Summary and Conclusions

7.1 Justification of Methodology for Estimating SMI and AMI Among Prisoners

As discussed in Chapter 6, there are several potential limitations of the methodology used to estimate SMI and AMI among federal and state prisoners in the 2016 SPI. In addition to the potential limitations associated with the general model-based approach to estimation, three limitations specific to this study arise from: (1) the *small sample size* of the CJ subsample of the 2008-2012 MHSS within which the SMI prediction model was developed, (2) the *limited information* in the 2016 SPI available for the development of SMI prediction models, and (3) the *different populations* represented by criminal justice status (i.e., parolees, probationers, and arrestees) among the general population versus federal and state prisoners.

Despite the potential limitations of the estimation methodology, some justification for its use is provided. First, the general methodology has been shown to be a rigorous approach that provides reasonably accurate estimates of SMI and AMI (both overall and at the domain level) given a sufficient sample size (CBHSQ, 2014a). If the same general approach is used on a subsample where the sample size is small (as in this study), it does not negate the rigor of the approach itself, but rather means the resulting estimates will be less accurate (i.e., they will have larger standard errors or confidence intervals, and may possess some undetected bias).

Second, even though the 2016 SPI does not collect all of the information required for the 2012 NSDUH prediction model (i.e., past year K6 scale, WHODAS impairment measure, past year suicidality, or past year MDE), it does collect data on the past month K6 scale, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety (which ended up being used for the SMI model K2DA). Chapters 4 and 5 demonstrate that there is an association between SMI and AMI, and between past month K6 scale, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety. Thus, estimates resulting from SMI model K2DA will have higher associated TCE rates (i.e., sum of false positive and false negative rates in the ROC analysis), but they will still be (nearly) unbiased.

Third, even if the two populations (i.e., CJ population versus prisoner population) used to obtain and apply the estimates differ, it does not automatically mean that a prediction model that was fit to one population will not apply to a different population. For example, SMI model K2DA is based on past month K6 scores, lifetime self-reported diagnosed depression, and lifetime self-reported diagnosed anxiety, and these terms are assumed to apply similarly to both populations. The higher estimates observed in the prisoner population reflect that this is a population with higher levels of psychological distress, depression, and anxiety (Bronson & Berzofsky, 2017), which is exactly what the model specifies. In addition, the patterns of estimates of SMI and AMI among domains such as sex and race/Hispanic origin are consistent between the two populations and with those reported in three prior BJS reports (Bronson & Berzofsky, 2017; James & Glaze, 2006; Ditton, 1999). Patterns of estimates for different age groups are consistent with those reported in Bronson and Berzofsky (2017), and patterns of estimates for federal and state prisoners are consistent with those reported in

James and Glaze (2006) and Ditton (1999). This approach to estimating SMI and AMI among the prisoner population could also be used in future SPI surveys to track trends.

Therefore, giving due consideration to the limitations of the methodology, the various analyses suggest that SMI model K2DA, with appropriate cut points, provides estimates of mental illness among the prisoner population represented by the SPI that are in line with expectations and represent standard measures (SMI and AMI) used by other federal agencies in policymaking and research.

7.2 Guidelines on the Use of the SMI and AMI Estimation Methodology

Considering the strengths of the estimation methodology that have been demonstrated when the sample size is sufficiently large (e.g., CBHSQ, 2014a), this methodology has value in providing "baseline" estimates of SMI and AMI among prisoners. However, also considering the limitations associated with this methodology applied to the 2016 SPI data, these estimates should be used with some caution, particularly those at the domain level (e.g., sex or age group).

Examples of how the estimates could be appropriately used include—

- The estimates could be compared to those from the 2016 NSDUH to show how much higher estimates of SMI and AMI are among the prisoner population than among the general adult population not in prison. For example, the estimate of SMI among the general adult population in 2016 is 4.2% and 18.3% for AMI.[22] Among the prisoner population, the estimates are substantially higher (16.7% for SMI and 35.8% for AMI). Thus, even if the actual magnitude of the differences in estimates between the prisoner and general adult populations contains error, at least the differences appear to be in the expected direction. The higher rates of SMI and AMI estimated in the 2016 SPI may be a consequence of prisoners experiencing higher levels of psychological distress, depression, or anxiety. There is no way to test this assumption with the current information available, but it does appear to be consistent with the findings of Bronson and Berzofsky (2017) that rates of psychological distress, depression, or anxiety are higher in the prisoner population than among the general population.
- The estimate of SMI among federal prisoners is 7.6% and 18.0% for state prisoners. Again, even if these estimates contain some error, the difference between them closely mirrors the difference in estimates between federal and state prisoners using a different estimation methodology. For example, in Ditton (1999) estimates of mental illness[23] are 7.4% among federal prisoners and 16.2% for state prisoners. In addition, the direction of the difference in AMI estimates between federal and state prisoners is preserved. For example, estimates of AMI are 21.0% among federal prisoners and 37.9% for state prisoners.

[22] The 2008-2012 NSDUH estimate of SMI is 3.9% and 18.1% for AMI, thereby indicating that these estimates are fairly stable over time.
[23] In Ditton (1999), prisoners were identified as mentally ill if they met one of the following two criteria: (1) they reported a current mental or emotional condition; or (2) they reported an overnight stay in a mental hospital, unit, or treatment program during their lifetime.

- The estimate of SMI among male prisoners is 15.2% and 36.7% for female prisoners, and the estimate of AMI among male prisoners is 33.9% and 61.7% for female prisoners. Due to the smaller sample sizes associated within the two sexes, these estimates will be less accurate than the corresponding overall estimates and should be treated with caution. The direction of the difference for both SMI and AMI appears to be correct (CBHSQ, 2014a; Bronson & Berzofsky, 2017).
- The estimate of SMI among prisoners ages 18 to 24 is 15.4%, ages 25 to 34 is 17.4%, and age 35 or older is 16.6%. The estimate of AMI among prisoners ages 18 to 24 is 33.5%, ages 25 to 34 is 37.7%, and age 35 or older is 35.2%. Again, due to the smaller sample sizes associated within the three age groups, these estimates will likely be less accurate than the corresponding overall estimates and also should be treated with caution. However, the similarity of the estimates among the age groups may be credible because higher levels of psychological distress, depression, or anxiety in the prisoner population (Bronson & Berzofsky, 2017) may almost completely override the age-group differences observed in the general adult population.
- The estimate of SMI among non-Hispanic white prisoners is 22.9%, non-Hispanic black prisoners is 10.7%, and Hispanic prisoners is 13.6%. The estimate of AMI among non-Hispanic white prisoners is 46.0%, non-Hispanic black prisoners is 26.9%, and Hispanic prisoners is 29.7%. Due to the smaller sample sizes associated within the three race and Hispanic origin groups, these estimates will be less accurate than the corresponding overall estimates and should be treated with caution. The pattern of differences among non-Hispanic white, non-Hispanic black, and Hispanic prisoners appears to be consistent (CBHSQ, 2014a; Bronson & Berzofsky, 2017).

7.3 Future Research

If there is a future need to produce estimates of SMI and AMI that are more accurate than those provided in Table 19, BJS could consider the following changes for SPI:

- Include questions in future SPI questionnaires to capture information on the past year K6 scale, WHODAS impairment measure, past year suicidality, and past year MDE.
- Select a subsample of prisoners from a future SPI to undergo clinical interviews to provide gold-standard measures of SMI and AMI. The sample size of this subsample should be sufficiently large to provide reasonably accurate estimates overall and at the domain level.
- Conduct a study to attempt to replicate the results from the project on which BJS collaborated with NIMH and Dr. Trestman to develop and validate a scale to directly measure SMI and specific mental disorders through SPI (and other BJS inmate surveys). If the results can be replicated, then include that scale in future SPI questionnaires.

Table 19. Model-Based Estimates of SMI and AMI Among Prisoners, by Jurisdiction, Sex, Age, and Race/Hispanic Origin, 2016 SPI

	SMI	AMI
Total	16.7%	35.8%
Jurisdiction		
Federal	7.6%	21.0%
State	18.0%	37.9%
Sex		
Male	15.2%	33.9%
Female	36.7%	61.7%
Age		
18-24	15.4%	33.5%
25-34	17.4%	37.7%
35 or Older	16.6%	35.2%
Race/Hispanic Origin		
Non-Hispanic White	22.9%	46.0%
Non-Hispanic Black	10.7%	26.9%
Hispanic	13.6%	29.7%
Non-Hispanic Other	22.7%	45.5%

AMI = any mental illness; SMI = serious mental illness; SPI = Survey of Prison Inmates.

Note: Standard errors of estimates in this table are contained in Table B.19 in Appendix B.

Dataset = 2016 SPI data.
Analysis weight = WT_FINAL.

Source: Bureau of Justice Statistics, Survey of Prison Inmates, 2016.

References

Aldworth, J., Barnett-Walker, K., Chromy, J., Karg, R., Morton, K., Novak, S., & Spagnola, K. (2009). Measuring serious mental illness with the NSDUH: Results of 2008 12-month analysis. In Substance Abuse and Mental Health Services Administration, *2008 National Survey on Drug Use and Health: Methodological resource book*.

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

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

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR)* (4th ed., text revision).

Bronson, J., & Berzofsky, M. (2017). *Indicators of mental health problems reported by prisoners and jail inmates, 2011-12* (NCJ 250612). Bureau of Justice Statistics.

Carson, E. A. (2014). *Prisoners in 2013* (NCJ 247282). Bureau of Justice Statistics.

Center for Behavioral Health Statistics and Quality. (2014a). *2012 National Survey on Drug Use and Health: Methodological resource book section 16a: 2012 Mental Health Surveillance Study design and estimation report.* Substance Abuse and Mental Health Services Administration.

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

Center for Behavioral Health Statistics and Quality. (2015a). *2012 National Survey on Drug Use and Health: Estimating mental illness among adults in the United States: Revisions to 2008 estimation procedures*. Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2015b). *2014 National Survey on Drug Use and Health: Methodological summary and definitions*. Substance Abuse and Mental Health Services Administration.

Center for Behavioral Health Statistics and Quality. (2017). *2015 National Survey on Drug Use and Health: Methodological resource book section 13: Statistical inference report.* Substance Abuse and Mental Health Services Administration.

Colpe, L. J., Barker, P. R., Karg, R. S., Batts, K. R., Morton, K. B., Gfroerer, J. C., Stolzenberg, S. J., Cunningham, D. B., First, M. B., & Aldworth, J. (2010). The National Survey on Drug Use and Health Mental Health Surveillance Study: Calibration study design and field procedures. *International Journal of Methods in Psychiatric Research, 19*(S1), 36-48.

Crippa, J. A. S., de Lima Osório, F., Del-Ben, C. M., Santos Filho, A., da Silva Freitas, M. C., & Loureiro, S. R. (2008). Comparability between telephone and face-to-face structured clinical interview for *DSM-IV* in assessing social anxiety disorder. *Perspectives in Psychiatric Care, 44*(4), 241-247.

Ditton, P. M. (1999). *Mental health and treatment of inmates and probationers* (NCJ 174463). Bureau of Justice Statistics.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The global assessment scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33*(6), 766-771.

Feucht, T. E., & Gfroerer, J. (2011). *Mental and substance use disorders among adult men on probation or parole: Some success against a persistent challenge*. Substance Abuse and Mental Health Services Administration.

Fennig, S., Craig, T., Lavelle, J., Kovasznay, B., & Bromet, E. J. (1994). Best-estimate versus structured interview-based diagnosis in first-admission psychosis. *Comprehensive Psychiatry, 35*(5), 341-348.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, research version, non-patient edition* (SCID-I/NP). New York State Psychiatric Institute.

Glasheen, C., Hedden, S. L., Kroutil, L. A., Pemberton, M. R., & Godstrom, I. (2012). *Past year arrest among adults in the United States: Characteristics of and association with mental illness and substance use*. Substance Abuse and Mental Health Services Administration.

Hajebi, A., Motevalian, A., Amin-Esmaeili, M., Hefazi, M., Radgoodarzi, R., Rahimi-Movaghar, A., & Sharifi, V. (2012). Telephone versus face-to-face administration of the structured clinical interview for *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, for diagnosis of psychotic disorders. *Comprehensive Psychiatry, 53*(5), 579-583.

Hedden, S., Gfroerer, J., Barker, P., Smith, S., Pemberton, M. R., Saavedra, L. M., Forman-Hoffman, V. L., Ringeisen, H., & Novak, S. P. (2012). *Comparison of NSDUH mental health data and methods with other data sources.* Substance Abuse and Mental Health Services Administration.

Jackson, J. S., Neighbors, H. W., Nesse, R. M., Trierweller, S. J., & Torres, M. (2004). Methodological innovations in the National Survey of American Life. *International Journal of Methods in Psychiatric Research, 13*(4), 289-298.

James, D. J., & Glaze, L. E. (2006). *Mental health problems of prison and jail inmates* (NCJ 213600). Bureau of Justice Statistics.

Jordan, B. K., Karg, R. S., Batts, K. R., Epstein, J. F., & Wiesen, C. (2008). A clinical validation of the National Survey on Drug Use and Health assessment of substance use disorders. *Addictive Behaviors*, *33*(6), 782-798.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). A population based twin study of major depression in women: The impact of varying definitions of illness. *Archives of General Psychiatry, 49*(4), 257-266.

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

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*(2), 184-189.

Kranzler, H. R., Kadden, R. M., Babor, T. F., Tennen, H., & Rounsaville, B. J. (1996). Validity of the SCID in substance abuse patients. *Addiction, 91*(6), 859-868.

Kranzler, H. R., Kadden, R. M., Burleson, J. A., Babor, T. F., Apter, A., & Rounsaville, B. J. (1995). Validity of psychiatric diagnoses in patients with substance use disorders: Is the interview more important than the interviewer? *Comprehensive Psychiatry, 36*(4), 278-288.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lattimore, P. K., Steffey, D. M., Gfroerer, J., Bose, J., Pemberton, M. R., & Penne, M. A. (2014). *Arrestee substance use: Comparison of estimates from the National Survey on Drug Use and Health and the Arrestee Drug Abuse Monitoring program*. Substance Abuse and Mental Health Services Administration.

Lee, S., Tsang, A., Lau, L., Mak, A., Ng, K. L., & Chan, D. M.-I. (2008). Concordance between telephone survey classification and face-to-face Structured Clinical Interview in the diagnosis of generalized anxiety disorder in Hong Kong. *Journal of Anxiety Disorders, 22*(8), 1403-1411.

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

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*(S1), 49-60.

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

Rehm, J., Üstün, T. B., Saxena, S., Nelson, C. B., Chatterji, S., Ivis, F., & Adlaf, E. (1999). On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. *International Journal of Methods in Psychiatric Research, 8*(2), 110-122.

Rohde, P., Lewinsohn, P. M., & Seeley, J. R. (1997). Comparability of telephone and face-to-face interviews in assessing axis I and axis II disorders. *The American Journal of Psychiatry, 154*(11), 1593-1598.

RTI International. (2012). *SUDAAN language manual, volumes 1 and 2, release 11.0.0*.

Segal, D. L., Kabacoff, R. I., Hersen, M., Van Hasselt, V. B., & Ryan, C. F. (1995). Update on the reliability of diagnosis in older psychiatric outpatients using the structured clinical interview for DSM-III-R. *Journal of Clinical Geropsychology, 1*(4), 313-321.

Shear, M. K., Greeno, C., Kang, J., Ludewig, D., Frank, E., Swartz, H. A., & Hanekamp, M. (2000). Diagnosis of nonpsychotic patients in community clinics. *The American Journal of Psychiatry, 157*(4), 581-587.

Sobin, C., Weissman, M. M., Goldstein, R. B., Adams, P., Wickramaratne, P., Warner, V., & Lish, J. D. (1993). Diagnostic interviewing for family studies: Comparing telephone and face-to-face methods for the diagnosis of lifetime psychiatric disorders. *Psychiatric Genetics, 3*(4), 227-233.

Steiner, J. L., Tebes, J. K., Sledge, W. H., & Walker, M. L. (1995). A comparison of the structured clinical interview for DSM-III-R and clinical diagnoses. *The Journal of Nervous and Mental Disease, 183*(6), 365-369.

Substance Abuse and Mental Health Services Administration. (2009). *Results from the 2008 National Survey on Drug Use and Health: National findings* (HHS Publication No. SMA 09-4434). U.S. Government Publishing Office.

Trestman, R.& Shelton, D. (2010). Mental health subscale development for national surveys of prisoners and jail inmates. Unpublished final report delivered to the Bureau of Justice Statistics through the National Institute of Mental Health, Research Infastructure Support Program, Grant No. 5R24MH067030-05.

Zanarini, M. C., & Frankenburg, F. R. (2001). Attainment and maintenance of reliability of axis I and II disorders over the course of a longitudinal study. *Comprehensive Psychiatry, 42*(5), 369-374.

Zanarini, M. C., Skodol, A. E., Bender, D. S., Dolan, R. T., Sanislow, C. A., Schaefer, E., Morey, L. C., Grilo, C. M., Shea, M. T., McGlashan, T. H., & Gunderson, J. G. (2000). The Collaborative Longitudinal Personality Disorders Study: Reliability of axis I and II diagnoses. *Journal of Personality Disorders, 14*(4), 291-299.

Appendix A: 2015 NSDUH Precision-Based Suppression Rules

The Center for Behavioral Health Statistics and Quality (CBHSQ; 2017) details the 2015 National Survey on Drug Use and Health (NSDUH) precision-based suppression rules. For convenience, this appendix displays the rules.

Suppression of Estimates with Low Precision in 2015 NSDUH

Direct survey estimates that were considered to be unreliable because of unacceptably large sampling errors were not reported, but rather were noted by an asterisk (*). *The criteria used to assess the need to suppress direct survey estimates were based on prevalence (for proportion estimates), the relative standard error (RSE) (defined as the ratio of the standard error [SE] over the estimate), nominal (actual) sample size, and effective sample size for each estimate.*

Proportion estimates ($\hat{p}$) within the range $0 < \hat{p} < 1$, and corresponding estimated numbers of users, were suppressed if—

$$\text{RSE}[-\ln(\hat{p})] > .175 \text{ when } \hat{p} \leq .5$$

or

$$\text{RSE}[-\ln(1-\hat{p})] > .175 \text{ when } \hat{p} > .5 .$$

The choice of .175 is arbitrary, but it roughly marks the tails of the distribution.

Based on a first-order Taylor series approximation of $\text{RSE}[-\ln(\hat{p})]$ and $\text{RSE}[-\ln(1-\hat{p})]$, the following equation was derived and used for computational purposes when applying a suppression rule dependent on effective sample sizes:

$$\frac{\text{SE}(\hat{p})/\hat{p}}{-\ln(\hat{p})} > .175 \text{ when } \hat{p} \leq .5$$

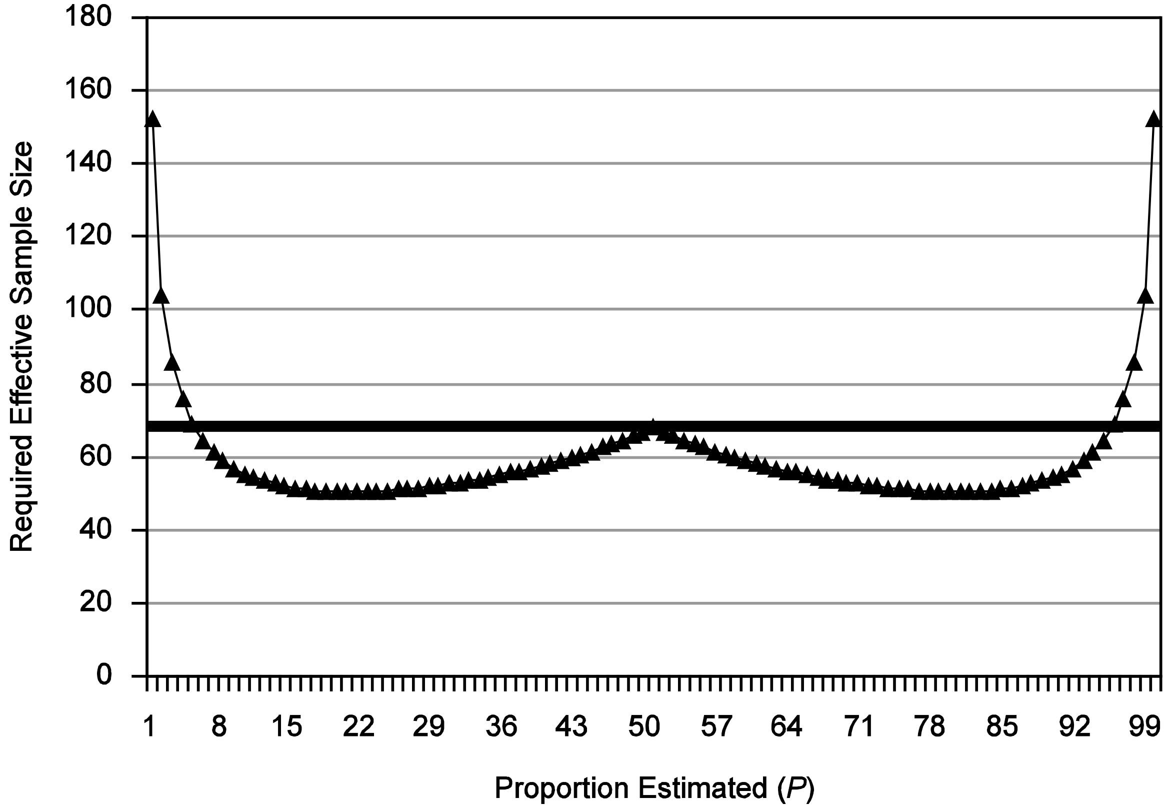
or

$$\frac{\text{SE}(\hat{p})\ /(1-\hat{p})}{-\ln(1-\hat{p})} > .175 \text{ when } \hat{p} > .5 .$$

The separate formulas for $\hat{p} \leq .5$ and $\hat{p} > .5$ produce a symmetric suppression rule; that is, if $\hat{p}$ is suppressed, $1-\hat{p}$ will be suppressed as well. See Figure A.1 for a graphical representation of the required minimum effective sample sizes as a function of the proportion estimated. When $.05 < \hat{p} < .95$, the symmetric properties of the rule produce local minimum effective sample sizes at

$\hat{p}$ = .2 and again at $\hat{p}$ = .8, such that an effective sample size of greater than 50 is required; this means that estimates would be suppressed for these values of $\hat{p}$ unless the effective sample sizes were greater than 50. Within this same interval of $.05 < \hat{p} < .95$, a local maximum effective sample size of 68 is required at $\hat{p}$ = .5.

Figure A.1. Required Effective Sample in the 2015 NSDUH as a Function of the Proportion Estimated



NSDUH = National Survey on Drug Use and Health.

These varying effective sample size requirements sometimes produced unusual occurences of suppression for a particular combination of prevalence estimates. For example, in some cases, lifetime prevalence estimates near $\hat{p}$ = .5 were suppressed (effective sample size was less than 68 but greater than 50), while not suppressing the corresponding past year or past month estimates near $\hat{p}$ = .2 (effective sample sizes greater than 50). *To reduce the occurrence of this type of inconsistency and maintain a conservative suppression rule, estimates of* $\hat{p}$ *between .05 and .95, which had effective sample sizes below 68, were suppressed starting with the 2000 NSDUH.*

The effective sample size for a domain is a function of the nominal sample size and the design effect (i.e., nominal sample size/design effect). During the original development of this suppression rule, the design effect was calculated outside SUDAAN® (RTI International, 2012) in SAS®. Since the 2005 NSDUH analysis, the direct SUDAAN design effect was used to provide a more precise and accurate reflection of the design effect (because of the removal of several possible rounding errors) when compared with the SAS method used in the past. The differences between the direct SUDAAN design effects and the SAS-calculated design effects occur only at approximately the tenth decimal place or later; however, previously published estimates that were on the borderline of being suppressed or unsuppressed because of the effective sample size suppression rule may potentially change from suppressed to unsuppressed, or vice versa.

Design effects range widely among the measures and domains found in the detailed tables. Potential problems with suppression occur only if large design effects are combined with small domains. Large estimates of design effects when resulting from small sample sizes (variability of the variance estimate) should be suppressed on effective sample size alone, and the rule above achieves this. *But to protect against unreliable estimates caused by small design effects and small nominal sample sizes, a minimum nominal sample size suppression criterion (n = 100) was employed starting with the 2000 NSDUH.*

Table A.1 shows a formula for calculating design effects. *Prevalence estimates also were suppressed if they were close to 0 or 100 percent (i.e., if* $\hat{p} < .00005$ *or if* $\hat{p} > .99995$*).*

Table A.1. Summary of 2015 NSDUH Suppression Rules

Estimate	Suppress if—
Prevalence Estimate, $\hat{p}$, with Nominal Sample Size, *n*, and Design Effect, *deff* $\left(deff = \frac{n[\mathrm{SE}(\hat{p})]^2}{\hat{p}(1-\hat{p})} \right)$	(1) The estimated prevalence estimate, $\hat{p}$, is < 0.00005 or > 0.99995,[1] or (2) $\frac{\mathrm{SE}(\hat{p})\ /\ \hat{p}}{-\ln(\hat{p})} > 0.175$ when $\hat{p} \le 0.5$, or $\frac{\mathrm{SE}(\hat{p})\ /\ (1-\hat{p})}{-\ln(1-\hat{p})} > .175$ when $\hat{p} > 0.5$, or (3) Effective *n* < 68, where $Effective\ n = \frac{n}{deff} = \frac{\hat{p}(1-\hat{p})}{[\mathrm{SE}(\hat{p})]^2}$, or (4) *n* < 100. Note: The rounding portion of this suppression rule for prevalence estimates will produce some estimates that round at one decimal place to 0.0 or 100.0 percent but are not suppressed from the tables.[2]
Estimated Number (Numerator of $\hat{p}$)	The estimated prevalence estimate, $\hat{p}$, is suppressed. Note: In some instances when $\hat{p}$ is not suppressed, the estimated number may appear as a 0 in the tables. This means that the estimate is greater than 0 but less than 500 (estimated numbers are shown in thousands). Note: In some instances when totals corresponding to several different means that are displayed in the same table, and some, but not all, of those means are suppressed, the totals will not be suppressed. When all means are suppressed, the totals will also be suppressed.
Means not Bounded Between 0 and 1 (i.e., Mean Age at First Use, Mean Number of Drinks), $\bar{x}$, with Nominal Sample Size, *n*	(1) $\mathrm{RSE}(\bar{x}) > 0.5$, or (2) *n* < 10.

deff = design effect; NSDUH = National Survey on Drug Use and Health; RSE = relative standard error; SE = standard error.

[1]Starting with the 2015 NSDUH, the close to 100 percent portion of the rule was changed to $\hat{p}$ > 0.99995, instead of the old rule, which was greater than or equal to 0.99995. This was done so the close to 0 and close to 100 rule were both strict inequalities.

[2]See Sections 3 and 7 of this report for more information on rounding.

Note: The suppression rules included in this table are used for detecting unreliable estimates and are sufficient for confidentiality purposes in the context of the first findings reports and detailed tables.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2015.

Beginning with the 1991 survey, the suppression rule for proportions based on $\mathrm{RSE}[-\ln(\hat{p})]$ described above replaced an older rule in which data were suppressed whenever RSE($\hat{p}$) > .5. This rule was changed because the older rule imposed a stringent application for small $\hat{p}$ but a very lax application for large $\hat{p}$. The new rule ensured a more uniformly stringent application across the whole range of $\hat{p}$ (i.e., from 0 to 1). The old rule also was asymmetric in the sense that suppression only occurred in terms of $\hat{p}$; that is, there was no complementary rule for (1 – $\hat{p}$), which the new suppression rules now account.

Estimates of totals were suppressed if the corresponding prevalence estimates were suppressed. Estimates of means not bounded between 0 and 1 (e.g., mean age at first use, mean number of drinks consumed) were suppressed if the RSEs of the estimates were larger than .5 or if the sample sizes were smaller than 10 respondents. This rule was based on an empirical examination of the estimates of mean age of first use and their SEs for various empirical sample sizes. *Although arbitrary, a sample size of 10 appears to provide sufficient precision and still allow reporting by year of first use for many substances.* In these cases, the totals (e.g., total number of drinks consumed) were suppressed if the corresponding mean estimates were suppressed.

Appendix B: Standard Errors of Estimates

This appendix contains standard errors of estimates of tables in the main report where space constraints did not allow the standard errors to appear. The numbering of the tables in this appendix corresponds with that of the original table and is prefixed with “B”. For example, if Table 10 in the main text did not include standard errors, then those standard errors are included in Table B.10.

Table B.10. Standard Errors of SMI Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH

Model		**K2A30**		**K2A1845**		**K2A2545**		**K2A1850**		**K2DA**		**K2**	
Dataset	**MHSS**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**
Estimator	**SCID**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**
Sample Size	**316**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**
Demographic Variable													
Total	3.29	2.89	0.60	2.88	0.60	2.88	0.58	2.89	0.59	3.10	0.54	3.04	0.52
Sex													
Male	3.46	3.11	0.65	3.12	0.64	3.12	0.62	3.13	0.62	3.36	0.55	3.09	0.57
Female	*7.08*	*6.21*	1.24	*6.10*	1.23	*6.10*	1.20	*6.10*	1.22	*6.31*	1.20	*7.40*	1.09
Age													
18-24	2.61	3.00	0.42	2.90	0.36	2.90	0.38	3.03	0.36	3.12	0.44	*5.61*	0.45
25-34	*5.85*	*4.27*	0.90	*4.27*	0.89	*4.27*	0.87	*4.27*	0.89	*6.11*	0.94	*6.21*	0.86
35 or Older	*6.06*	*5.70*	1.28	*5.70*	1.26	*5.70*	1.24	*5.70*	1.25	*5.45*	1.08	*4.57*	1.03
Race/Hispanic Origin													
Non-Hispanic White	4.41	3.75	0.88	3.73	0.88	3.73	0.85	3.73	0.86	4.19	0.77	3.74	0.65
Non-Hispanic Black	*5.77*	*10.76*	0.99	*10.76*	0.92	*10.76*	0.92	*10.76*	0.92	*10.76*	0.90	*8.14*	1.15
Hispanic	*1.75*	*2.13*	1.07	*2.06*	1.01	*2.06*	0.95	*2.06*	0.99	*2.07*	0.95	*11.18*	1.24
Non-Hispanic Other	*15.79*	*3.16*	2.32	*3.16*	2.33	*3.16*	2.30	*2.70*	2.30	*3.16*	2.41	*2.41*	2.00
Education													
Less than High School	*6.12*	*3.09*	0.93	*3.09*	0.88	*3.09*	0.88	*3.09*	0.88	*3.09*	0.83	*6.08*	0.93
High School Graduate	4.07	*5.85*	0.90	*5.84*	0.89	*5.84*	0.86	*5.86*	0.87	*5.67*	0.82	*5.53*	0.85
Some College	*4.75*	*4.34*	1.25	*4.34*	1.25	*4.34*	1.24	*4.34*	1.24	*5.75*	1.18	*4.11*	0.81
College Graduate	*7.32*	*2.57*	2.78	*2.57*	2.75	*2.57*	2.75	*2.57*	2.76	*2.57*	2.21	*1.38*	1.85
Marital Status													
Married	*7.50*	*4.69*	1.43	*4.69*	1.40	*4.69*	1.37	*4.69*	1.38	*3.95*	1.07	*5.82*	1.33
Widowed	*1.67*	*2.01*	*6.19*	*2.01*	*5.87*	*2.01*	*5.82*	*2.01*	*5.87*	*2.01*	4.74	*1.67*	*4.75*
Divorced or Separated	*12.66*	*16.66*	1.77	*16.64*	1.76	*16.64*	1.74	*16.64*	1.74	*16.63*	1.60	*11.31*	1.36
Never Married	3.07	2.66	0.58	2.62	0.57	2.62	0.56	2.67	0.57	3.34	0.56	3.90	0.53

(continued)

Table B.10. Standard Errors of SMI Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH (continued)

Model		**K2A30**		**K2A1845**		**K2A2545**		**K2A1850**		**K2DA**		**K2**	
Dataset	**MHSS**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**
Estimator	**SCID**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**	**Model**
Sample Size	**316**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**	**316**	**15,400**
Demographic Variable													
Military Service													
Yes	*7.01*	*7.53*	2.84	*7.53*	2.82	*7.53*	2.74	*7.53*	2.77	*7.53*	2.27	*11.72*	1.85
No	3.41	3.02	0.60	3.01	0.59	3.01	0.58	3.02	0.59	3.22	0.54	3.10	0.54
Received Mental Health Treatment in Past Year													
Yes	*10.97*	*11.03*	1.59	*10.97*	1.61	*10.97*	1.62	*10.97*	1.63	*10.38*	1.62	*10.13*	1.57
No	2.72	2.39	0.43	2.39	0.41	2.39	0.38	2.38	0.39	2.50	0.33	2.31	0.43
Illicit Drug or Alcohol Abuse or Dependence in Past Year													
Yes	*4.91*	*6.06*	1.00	*6.09*	0.99	*6.09*	0.97	*6.09*	0.98	*5.79*	0.91	*5.77*	0.93
No	*4.64*	*3.05*	0.73	3.04	0.71	3.04	0.70	3.02	0.70	3.47	0.63	3.79	0.57

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. The standard errors in this table correspond with the estimates contained in Table 10 in Chapter 4.

Predictor variables included in the following models:

K2A30: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE30
K2A1845: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1845
K2A2545: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE2545
K2A1850: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1850
K2DA: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U
K2: K6SCMON2

Datasets = CJ subsample of 2008-2012 adult NSDUH data for NSDUH estimates, and CJ subsample of 2008-2012 MHSS clinical data for MHSS estimates.
Analysis weights = ANALWT for NSDUH estimates, and MHFNLWGT for MHSS estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table B.11. Standard Errors of SMI Hybrid Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH

Model		**K2A30**	**K2A1845**	**K2A2545**	**K2A1850**	**K2DA**	**K2**
Original Cut Point		**0.24589**	**0.26490**	**0.28900**	**0.27548**	**0.25776**	**0.21262**
Hybrid Cut Point		**0.29126**	**0.30852**	**0.30971**	**0.29772**	**0.27879**	**0.21262**
Dataset	**MHSS**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**
Estimator	**SCID**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**
Sample Size	**316**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**
Demographic Variable							
Total	3.29	0.56	0.58	0.57	0.58	0.57	0.52
Sex							
Male	3.46	0.59	0.62	0.61	0.61	0.60	0.53
Female	*7.08*	1.20	1.21	1.21	1.21	1.19	1.19
Age							
18-24	2.61	0.39	0.32	0.33	0.34	0.38	0.43
25-34	*5.85*	0.88	0.87	0.85	0.87	0.86	0.93
35 or Older	*6.06*	1.18	1.22	1.22	1.22	1.20	1.05
Race/Hispanic Origin							
Non-Hispanic White	4.41	0.81	0.85	0.85	0.84	0.83	0.75
Non-Hispanic Black	*5.77*	0.94	0.91	0.91	0.89	0.89	0.89
Hispanic	*1.75*	0.92	0.94	0.90	0.93	0.88	0.94
Non-Hispanic Other	*15.79*	2.27	2.26	2.26	2.26	2.26	2.13
Education							
Less than High School	*6.12*	0.84	0.87	0.84	0.86	0.83	0.81
High School Graduate	4.07	0.83	0.84	0.83	0.82	0.81	0.81
Some College	*4.75*	1.19	1.24	1.24	1.23	1.21	1.12
College Graduate	*7.32*	2.62	2.75	2.75	2.74	2.74	2.20
Marital Status							
Married	*7.50*	1.26	1.35	1.34	1.29	1.27	1.04
Widowed	*1.67*	*5.44*	*5.63*	*5.73*	*5.81*	5.82	*4.74*
Divorced or Separated	*12.66*	1.70	1.73	1.73	1.73	1.73	1.57
Never Married	3.07	0.54	0.54	0.54	0.53	0.53	0.55

(continued)

Table B.11. Standard Errors of SMI Hybrid Cut Point Estimates of Selected Models for Adult Criminal Justice Population, 2008-2012 NSDUH (continued)

Model		**K2A30**	**K2A1845**	**K2A2545**	**K2A1850**	**K2DA**	**K2**
Original Cut Point		**0.24589**	**0.26490**	**0.28900**	**0.27548**	**0.25776**	**0.21262**
Hybrid Cut Point		**0.29126**	**0.30852**	**0.30971**	**0.29772**	**0.27879**	**0.21262**
Dataset	**MHSS**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**	**NSDUH**
Estimator	**SCID**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**	**Hybrid**
Sample Size	**316**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**	**15,400**
Demographic Variable							
Military Service							
Yes	*7.01*	2.70	2.72	2.69	2.75	2.73	2.22
No	3.41	0.56	0.58	0.57	0.57	0.56	0.53
Received Mental Health Treatment in Past Year							
Yes	*10.97*	1.64	1.66	1.65	1.67	1.64	1.62
No	2.72	0.36	0.38	0.37	0.37	0.36	0.31
Illicit Drug or Alcohol Abuse or Dependence in Past Year							
Yes	*4.91*	0.94	0.97	0.96	0.96	0.95	0.90
No	4.64	0.66	0.69	0.69	0.68	0.67	0.61

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. The standard errors in this table correspond with the estimates contained in Table 11 in Chapter 4. Original cut point derived to yield (nearly) unbiased cut point estimator in MHSS, and hybrid cut point derived to yield (nearly) unbiased cut point estimator in NSDUH.

Predictor variables included in the following models:

K2A30: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE30
K2A1845: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1845
K2A2545: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE2545
K2A1850: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U + AGE1850
K2DA: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U
K2: K6SCMON2

Datasets = CJ subsample of 2008-2012 adult NSDUH data for NSDUH estimates, and CJ subsample of 2008-2012 MHSS clinical data for MHSS estimates.
Analysis weights = ANALWT for NSDUH estimates, and MHFNLWGT for MHSS estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table B.12. Standard Errors of SMI Hybrid Cut Point Estimates of Model K2DA for Criminal Justice Subpopulations, 2008-2012 NSDUH

Criminal Justice Subpopulation	All		Parolees		Probationers		Parolees or Probationers	
Dataset	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**	**MHSS**	**NSDUH**
Estimator	**SCID**	**Hybrid**	**SCID**	**Hybrid**	**SCID**	**Hybrid**	**SCID**	**Hybrid**
Hybrid Cut Point		**0.27879**		**0.25525**		**0.27879**		**0.27659**
Sample Size	**316**	**15,400**	**56**	**2,300**	**186**	**8,300**	**195**	**8,800**
Demographic Variable								
Total	3.29	0.52	*5.76*	1.14	*5.00*	0.70	*4.62*	0.67
Sex								
Male	3.46	0.53	*6.05*	1.02	*4.69*	0.70	*4.74*	0.66
Female	*7.08*	1.19	*12.54*	3.71	*10.28*	1.54	*9.77*	1.53
Age								
18-24	2.61	0.43	*0.54*	1.44	*4.77*	0.59	*4.53*	0.58
25-34	*5.85*	0.93	*16.93*	2.06	*7.80*	1.23	*7.67*	1.15
35 or Older	*6.06*	1.05	*7.34*	1.89	*9.65*	1.42	*8.80*	1.31
Race/Hispanic Origin								
Non-Hispanic White	4.41	0.75	*6.91*	1.73	*8.14*	1.01	*7.50*	0.97
Non-Hispanic Black	*5.77*	0.89	*18.80*	2.45	*4.33*	1.39	*10.82*	1.24
Hispanic	*1.75*	0.94	*0.17*	1.82	*2.22*	1.06	*2.09*	1.06
Non-Hispanic Other	*15.79*	2.13	*18.70*	*6.03*	*5.23*	3.25	*5.22*	3.21
Education								
Less than High School	*6.12*	0.81	*17.24*	1.75	*7.68*	1.08	*8.88*	1.02
High School Graduate	4.07	0.81	*3.34*	2.04	*5.82*	1.04	*4.84*	1.00
Some College	*4.75*	1.12	*7.78*	2.39	*8.48*	1.56	*8.12*	1.50
College Graduate	*7.32*	2.20	*25.49*	*4.40*	*11.57*	2.68	*11.46*	2.54
Marital Status								
Married	*7.50*	1.04	*2.99*	1.94	*9.66*	1.32	*7.60*	1.26
Widowed	*1.67*	4.74	*	*12.36*	*	*8.05*	*	*7.45*
Divorced or Separated	*12.66*	1.57	*18.75*	3.46	*20.56*	2.21	*20.38*	2.05
Never Married	3.07	0.55	*7.51*	1.21	*4.92*	0.74	*4.81*	0.71

(continued)

Table B.12. Standard Errors of SMI Hybrid Cut Point Estimates of Model K2DA for Criminal Justice Subpopulations, 2008-2012 NSDUH (continued)

Criminal Justice Subpopulation	All		Parolees		Probationers		Parolees or Probationers	
Dataset	MHSS	NSDUH	MHSS	NSDUH	MHSS	NSDUH	MHSS	NSDUH
Estimator	SCID	Hybrid	SCID	Hybrid	SCID	Hybrid	SCID	Hybrid
Hybrid Cut Point		0.27879		0.25525		0.27879		0.27659
Sample Size	316	15,400	56	2,300	186	8,300	195	8,800
Demographic Variable								
Military Service								
Yes	*7.01*	2.22	*0.00*	*2.93*	*13.39*	2.35	*13.39*	2.19
No	3.41	0.53	*6.12*	1.21	*5.17*	0.73	*4.77*	0.69
Received Mental Health Treatment in Past Year								
Yes	*10.97*	1.62	*	3.79	*11.55*	2.10	*10.98*	2.02
No	2.72	0.31	*3.71*	0.79	*2.68*	0.41	*2.40*	0.41
Illicit Drug or Alcohol Abuse or Dependence in Past Year								
Yes	*4.91*	0.90	*11.65*	2.14	*8.19*	1.19	*8.18*	1.14
No	4.64	0.61	*6.33*	1.31	*6.10*	0.82	*5.38*	0.77

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness.

*Indicates actual suppression due to disclosure risk from small denominators.

Note: Due to standard CBHSQ disclosure limitation protocols, certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. The standard errors in this table correspond with the estimates contained in Table 12 in Chapter 4. Hybrid cut point was derived to yield (nearly) unbiased cut point estimator in NSDUH.

Predictor variables included in the following model:

K2DA: K6SCMON2 + DEPRSLIF_U + ANXDLIF_U

Datasets = CJ subsample of 2008-2012 adult NSDUH data for NSDUH estimates, and CJ subsample of 2008-2012 MHSS clinical data for MHSS estimates.

Analysis weights = ANALWT for NSDUH estimates, and MHFNLWGT for MHSS estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, 2008-2012.

Table B.13. Standard Errors of SMI Estimates Based on Model K2DA for Different Data, Cut Points, and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI

Dataset	2008-2012 MHSS			2008-2012 NSDUH			2016 SPI			
Subpopulation	Criminal Justice		Parolees or Probationers	Criminal Justice		Parolees or Probationers	All Prisoners			
Estimator	SCID	MHSS Cut Point	SCID	MHSS Cut Point	CJ Hybrid Cut Point	PP Hybrid Cut Point	MHSS Cut Point	CJ Hybrid Cut Point	PP Hybrid Cut Point	SPI Hybrid Cut point
Cut Point		0.25776		0.25776	0.27879	0.27659	0.25776	0.27879	0.27659	0.37446
Sample Size	316	316	195	15,400	15,400	8,800	24,848	24,848	24,848	24,848
Demographic Variable										
Total	3.29	3.10	*4.62*	0.54	0.52	0.67	0.53	0.52	0.52	0.48
Sex										
Male	3.46	3.36	*4.74*	0.55	0.53	0.66	0.56	0.55	0.56	0.51
Female	*7.08*	*6.31*	*9.77*	1.20	1.19	1.53	1.18	1.17	1.17	1.10
Age										
18-24	2.61	3.12	*4.53*	0.44	0.43	0.58	1.27	1.21	1.21	1.12
25-34	*5.85*	*6.11*	*7.67*	0.94	0.93	1.15	0.77	0.76	0.77	0.68
35 or Older	*6.06*	*5.45*	*8.80*	1.08	1.05	1.31	0.60	0.59	0.59	0.56
Race/Hispanic Origin										
Non-Hispanic White	4.41	4.19	*7.50*	0.77	0.75	0.97	0.79	0.77	0.77	0.78
Non-Hispanic Black	*5.77*	*10.76*	*10.82*	0.90	0.89	1.24	0.62	0.60	0.60	0.51
Hispanic	*1.75*	*2.07*	*2.09*	0.95	0.94	1.06	0.98	0.97	0.97	0.86
Non-Hispanic Other	*15.79*	*3.16*	*5.22*	2.41	2.13	3.21	1.08	1.08	1.08	1.03
Marital Status										
Married	*7.50*	*3.95*	*7.60*	1.07	1.04	1.26	0.97	0.95	0.96	0.84
Widowed	*1.67*	*2.01*	*20.84*	4.74	4.74	*7.45*	1.82	1.81	1.81	1.78
Divorced or Separated	*12.66*	*16.63*	*20.38*	1.60	1.57	2.05	0.79	0.77	0.77	0.78
Never Married	3.07	3.34	*4.81*	0.56	0.55	0.71	0.57	0.56	0.56	0.50
Military Service										
Yes	*7.01*	*7.53*	*13.39*	2.27	2.22	2.19	1.43	1.44	1.44	1.23
No	3.41	3.22	*4.77*	0.54	0.53	0.69	0.54	0.53	0.54	0.48

(continued)

Table B.13. Standard Errors of SMI Estimates Based on Model K2DA for Different Data, Cut Points, and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI (continued)

Dataset	**2008-2012 MHSS**			**2008-2012 NSDUH**			**2016 SPI**			
Subpopulation	**Criminal Justice**		**Parolees or Probationers**	**Criminal Justice**		**Parolees or Probationers**	**All Prisoners**			
Estimator	**SCID**	**MHSS Cut Point**	**SCID**	**MHSS Cut Point**	**CJ Hybrid Cut Point**	**PP Hybrid Cut Point**	**MHSS Cut Point**	**CJ Hybrid Cut Point**	**PP Hybrid Cut Point**	**SPI Hybrid Cut point**
Cut Point		**0.25776**		**0.25776**	**0.27879**	**0.27659**	**0.25776**	**0.27879**	**0.27659**	**0.37446**
Sample Size	**316**	**316**	**195**	**15,400**	**15,400**	**8,800**	**24,848**	**24,848**	**24,848**	**24,848**
Demographic Variable										
Jurisdiction										
Federal	~	~	~	~	~	~	0.81	0.73	0.75	0.61
State	~	~	~	~	~	~	0.59	0.58	0.59	0.54
Offense										
Violent	~	~	~	~	~	~	0.69	0.68	0.68	0.64
Property	~	~	~	~	~	~	0.97	0.98	0.98	0.96
Drug	~	~	~	~	~	~	0.63	0.61	0.62	0.56
Public Order	~	~	~	~	~	~	5.83	5.83	5.83	5.82
Other	~	~	~	~	~	~	:	:	:	:

CBHSQ = Center for Behavioral Health Statistics and Quality; CJ = criminal justice; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; PP = parolees or probationers; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness; SPI = Survey of Prison Inmates.

:Not calculated.

~Not applicable.

*Indicates actual suppression due to disclosure risk from small denominators.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. The standard errors in this table correspond with the estimates contained in Table 13 in Chapter 4.

Datasets = 2008-2012 MHSS clinical data for MHSS estimates, 2008-2012 adult NSDUH data for NSDUH estimates, and 2016 SPI data for SPI estimates.
Analysis weights = MHFNLWGT for MHSS estimates, ANALWT for NSDUH estimates, and WT_FINAL for SPI estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012; and Bureau of Justice Statistics, Survey of Prison Inmates, 2016.

Table B.17. Standard Errors of AMI Estimates Based on SMI Model K2DA for Different Data and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI

Dataset	2008-2012 MHSS		2008-2012 NSDUH	2016 SPI
Subpopulation	Criminal Justice		Criminal Justice	All Prisoners
Estimator	SCID	MHSS Cut Point	MHSS Cut Point	MHSS Cut Point
Cut Point		0.12505	0.12505	0.12505
Sample Size	316	316	15,400	24,848
Demographic Variable				
Total	4.62	4.35	0.70	0.71
Sex				
Male	5.08	4.77	0.76	0.76
Female	*8.12*	*9.98*	1.31	1.32
Age				
18-24	*6.45*	*5.26*	0.54	1.58
25-34	*6.62*	*8.46*	1.16	1.00
35 or Older	*8.54*	*7.91*	1.37	0.75
Race/Hispanic Origin				
Non-Hispanic White	*6.47*	*5.70*	0.94	0.94
Non-Hispanic Black	*10.20*	*11.77*	1.30	0.87
Hispanic	*5.73*	*12.71*	1.52	1.23
Non-Hispanic Other	*16.24*	*4.62*	2.66	1.32
Marital Status				
Married	*8.87*	*8.24*	1.66	1.24
Widowed	*24.08*	*2.01*	*6.75*	2.45
Divorced or Separated	*16.60*	*16.78*	1.90	1.02
Never Married	5.01	5.03	0.69	0.78
Military Service				
Yes	*16.16*	*19.06*	3.00	1.66
No	4.85	4.40	0.70	0.74

(continued)

Table B.17. Standard Errors of AMI Estimates Based on SMI Model K2DA for Different Data and Subpopulations, 2008-2012 MHSS, 2008-2012 NSDUH, and 2016 SPI (continued)

Dataset	2008-2012 MHSS		2008-2012 NSDUH	2016 SPI
Subpopulation	Criminal Justice		Criminal Justice	All Prisoners
Estimator	SCID	MHSS Cut Point	MHSS Cut Point	MHSS Cut Point
Cut Point		0.12505	0.12505	0.12505
Sample Size	316	316	15,400	24,848
Demographic Variable				
Jurisdiction				
Federal	~	~	~	1.29
State	~	~	~	0.79
Offense				
Violent	~	~	~	0.89
Property	~	~	~	1.16
Drug	~	~	~	0.85
Public Order	~	~	~	5.96
Other	~	~	~	:

AMI = any mental illness; CBHSQ = Center for Behavioral Health Statistics and Quality; MHSS = Mental Health Surveillance Study; NSDUH = National Survey on Drug Use and Health; SCID = Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition; SMI = serious mental illness; SPI = Survey of Prison Inmates.

:Not calculated.

~Not applicable.

Note: Due to standard CBHSQ disclosure limitation protocols certain unweighted NSDUH sample sizes were rounded to the nearest 100. Italicized estimate indicates estimate would be suppressed due to low precision according to NSDUH suppression rules. The standard errors in this table correspond with the estimates contained in Table 17 in Chapter 4.

Datasets = 2008-2012 MHSS clinical data for MHSS estimates, 2008-2012 adult NSDUH data for NSDUH estimates, and 2016 SPI data for SPI estimates.
Analysis weights = MHFNLWGT for MHSS estimates, ANALWT for NSDUH estimates, and WT_FINAL for SPI estimates.

Source: Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, National Survey on Drug Use and Health, Mental Health Surveillance Study clinical sample, 2008-2012; and Bureau of Justice Statistics, Survey of Prison Inmates, 2016.

Table B.19. Standard Errors of Model-Based Estimates of SMI and AMI Among Prisoners, by Jurisdiction, Sex, Age, and Race/Hispanic Origin, 2016 SPI

	SMI	AMI
Total	0.48	0.71
Jurisdiction		
Federal	0.61	1.29
State	0.54	0.79
Sex		
Male	0.51	0.76
Female	1.10	1.32
Age		
18-24	1.12	1.58
25-34	0.68	1.00
35 or Older	0.56	0.75
Race/Hispanic Origin		
Non-Hispanic White	0.78	0.94
Non-Hispanic Black	0.51	0.87
Hispanic	0.86	1.23
Non-Hispanic Other	1.03	1.32

AMI = any mental illness; SMI = serious mental illness; SPI = Survey of Prison Inmates.

Note: The standard errors in this table correspond with the estimates contained in Table 19 in Chapter 7.

Dataset = 2016 SPI data.
Analysis weight = WT_FINAL.

Source: Bureau of Justice Statistics, Survey of Prison Inmates, 2016.