# Development and Validation of a Brief Mental Health Screening Instrument for Newly Incarcerated Adults

**Julian D. Ford**
**Robert L. Trestman**
**Valerie Wiesbrock**
**Wanli Zhang**
*University of Connecticut School of Medicine*

*The authors report the development and initial psychometric evaluation of gender-specific brief screening instruments to identify undetected psychiatric impairment on incarceration. Women and men completed the Correctional Mental Health Screen (CMHS), a 56-item screen derived from validated measures. Representative subsamples completed structured diagnostic interviews within 5 days. An 8-item screen for women and a 12-item screen for men identified inmates with current Axis I psychiatric disorders with 83% to 100% accuracy on the basis of cut points chosen to maximize negative predictive power. The CMHS showed evidence of incremental predictive utility compared with two previously validated correctional mental health screening measures with White and Black men and White women. Incremental validity was not supported with Black women, for whom the CMHS performed well in identifying true cases but not in ruling out noncases. Analyses of internal consistency, interrater, and retest reliability and convergent, discriminant, and criterion validity supported the psychometric status of the CMHS.*

***Keywords:*** mental health; screening; adult; criminal justice; psychometrics

Psychiatric morbidity is epidemic but often not detected or treated in adult correctional institutions. Epidemiological studies with formal psychometric assessments in jails and prisons suggest a lifetime prevalence of psychiatric disorders ranging from 64% to 81% (Brinded, Simpson, Laidlaw, Fairley, & Malcolm, 2001; Diamond, Wang, Holzer, Thomas, & Cruser, 2001; Jordan, Schlenger, Fairbank, & Cadell, 1996; Teplin, 1994; Teplin, Abram, & McLelland, 1996) and current or past-6-month prevalence estimates nearly as high for women (48% to 71%; Jordan et al., 1996; Teplin et al., 1996) and half as high for men (31%; Gavin, Parsons, & Grubin, 2003; Teplin, 1994). Substantial lifetime prevalence estimates and levels of comorbidity (Abram, Teplin, & McLelland, 2003) among incarcerated adults have been reported for affective disorders (major depression 7% to 21%, dysthymia 4% to 14%, bipolar disorder 2% to 4%; Diamond et al., 2001; Jordan et al., 1996; Teplin et al., 1996), schizophrenia (1.5% to 5%; Diamond et al., 2001;

This research was funded solely through a grant from the National Institute of Justice (2000-IJ-CX-0044, Julian D. Ford and Robert L. Trestman, principal investigators). We gratefully acknowledge the collaboration and support of the Connecticut Department of Correction and the University of Connecticut Health Center's Correctional Managed Health Care; the research assistance of Karen Pagano, Susan Quarti, Jessie Wagner, and Kelly Beckom; and the technical assistance of Amy Stomsky. Correspondence concerning this article should be addressed to Julian D. Ford, University of Connecticut Health Center, Department of Psychiatry, MC1410, 263 Farmington Avenue, Farmington, CT 06030; e-mail: ford@psychiatry.uchc.edu.

*Assessment*, Volume 14, No. 3, September 2007 279-299
DOI: 10.1177/1073191107302944

Jordan et al., 1996; Parsons, Walker, & Grubin, 2001; Teplin, 1990, 1994; Teplin et al., 1996), generalized anxiety and panic disorder (2% to 8%; Jordan et al., 1996; Teplin, 1994; Teplin et al., 1996), post-traumatic stress disorder (PTSD) (33.5%; Teplin et al., 1996), and antisocial personality disorder (ASPD) (12% to 75%; Diamond et al., 2001; Jordan et al., 1996; Teplin et al., 1996).

Despite this evidence of prevalent mental health morbidity, fewer than one in four incarcerated adults with psychiatric disorders are identified in routine entry screening (Jordan et al., 2002; Parsons et al., 2001; Teplin, Abram, & McLelland, 1997). In the past 15 years, many state departments of correction have mandated mental health screening by health care professionals for all newly admitted inmates (Metzner, Miller, & Kleinsasser, 1994). By 2000, almost 70% of state prisons had formal policies mandating the screening of inmates at intake, and almost two thirds conducted psychiatric assessments (Beck & Maruschak, 2001). Yet mental health screening is uncommon at entry to jail (Teplin et al., 1997). Compared with prison inmates, jail detainees with mental illness are 50% less likely to receive mental health services and almost 200% less likely to receive counseling or therapy (Ditton, 1999). Mental illness in jails is a serious problem not just for detainees but also for the safety and effectiveness of custody procedures: Jailed adults with mental illness are 50% more likely than other jail inmates to have serious disciplinary problems (Ditton, 1999).

Several brief screening instruments have been developed to identify undetected and untreated psychiatric disorders in mental health treatment (Daradkeh, Ghubash, El-Rufaie, & Abou-Saleh, 1999; Zimmerman & Mattia, 2001), medical (Boutin-Foster, Ferrando, & Charlson, 2003; Ericsson et al., 2002; Furukawa, Goldberg, Rabe-Hesketh, & Uestuen, 2001; Herrmann, 1997; Spitzer, Kroenke, & Williams, 1999), and community (Kessler et al., 2002) populations. Screening measures assessing attitudes or personality characteristics that may lead to disciplinary problems (especially risk for violence) have been developed for adult corrections (Cooke, 1998; Walters & Chlumsky, 1993), but screening measures that target psychiatric disorders are less often reported. Relatively brief (i.e., 21- to 36-item) screening instruments have shown promise when evaluated psychometrically in adult prisons, although overidentification (i.e., false positives) is problematic (Andersen, Sestoft, Lillebaek, Gabrielson, & Hemmingsen, 2002; Book, Knap, & Holden, 2001; Boothby & Durham, 1999; Smith & Borland, 1999).

For adult jail populations, Teplin and Swartz (1989) statistically derived the 14-item Referral Decision Scale (RDS) from the Diagnostic Interview Schedule, with subscales for depression, bipolar disorder, and schizophrenia showing average sensitivity of 0.79 to 0.88 and average specificity of 0.99 for predicting full Diagnostic Interview Schedule diagnoses. Replication studies have shown the RDS to have strong negative predictive power (NPP) (>0.85) but only low (0.15 to 0.63) positive predictive power (PPP) in relation to structured interview (DiCataldo, Greer, & Profit; 1995; Hart, Roesch, Corrado, & Cox, 1993; McLearen & Ryba, 2003) or clinician diagnoses (Veysey, Steadman, Morrissey, Johansen, & Beckstead, 1998) and mixed evidence of convergent validity between the RDS and independent measures of bipolar disorder (mean $r = .19$), schizophrenia (mean $r = .29$), and depression (mean $r = .42$) (Rogers, Sewell, Ustad, Reinhardt, & Edwards 1995). With the exception of 10% of Veysey et al.'s (1998) sample, the RDS has not been evaluated with female jail detainees. Furthermore, the RDS does not address anxiety disorders or Axis II disorders.

A further refinement of the RDS, the 8-item Brief Jail Mental Health Screen (BJMHS; Steadman, Scott, Osher, Agnese, & Robbins, 2005), was designed to identify the presence of any of nine *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.) (*DSM-IV*; American Psychiatric Association, 1994) Axis I mood (e.g., major depression, bipolar disorder) or psychotic (e.g., schizophrenia, schizoaffective disorder) diagnoses. In a sample of newly admitted inmates who were selected to represent both individuals receiving referrals for mental health services and those who were not identified as warranting mental health referrals, the BJMHS correctly classified 74% of men and 62% of women who met criteria for index mood or psychotic disorders on the Structured Clinical Interview for *DSM-IV* (SCID). The BJMHS, like the RDS, also does not address anxiety or Axis II disorders, and although promising for men, had a high rate of false negatives (35%) with women. Furthermore, the SCID validation cohort oversampled inmates who were identified by the BJMHS as needing mental health services; it was not a representative sample of all inmates or of inmates with undetected mental health needs.

Thus, no brief mental health screening instrument has yet been determined to have robust clinical utility in identifying male and female jail detainees with undetected psychiatric morbidity. Therefore, in the current study, we used several brief but comprehensive self-report instruments (including the RDS) that were administered as a structured interview (the "composite screen") by research assessors (RAs) within the first 24 to 72 hours of detention in adult jails. The goal was to identify the briefest subset of items that achieved sensitive and specific prediction of current research interview–derived *DSM-IV* Axis I or II psychiatric disorders (excluding substance-use disorders) separately for men and women and for persons of different ethnocultural backgrounds. Item reduction procedures designed to eliminate low base rate and redundancy were conducted. A

subset of items with the best predictive utility was selected. Although psychological test scores may be temporally reliable over a 2-week period in the 1st month following admission to prison (Gagliardi, Lovell, Peterson, & Jemelka, 2004), new jail detainees experience substantial and variable stressors that may lead to poor retest reliability on screening tests. Therefore, we readministered the initial screen to a randomly selected sample of participants to determine if the results were sufficiently stable over the first few days in jail to constitute a reliable index of risk for psychiatric problems. To examine the predictive utility of the brief screening scales for men and women, we first conducted discriminant function analyses and then examined receiver-operating characteristic (ROC) analyses of sensitivity and specificity at different cut points and compared these results with those of parallel tests using items from two previously validated screening measures. Finally, the convergent and discriminant validity of the derived screens were tested in relation to archival corrections data.

The overall goal was to develop brief screening tools suitable to accomplish the first stage of the two-stage strategy articulated by Shrout, Skodol, and Dohrenwend (1986), that is, a low-cost, practical basis for selecting high-risk individuals who subsequently receive a Stage 2 intensive research or clinical assessment. By testing newly incarcerated men and women who were not identified in institutional intake screening procedures, we were able to extract items from a set of previously developed screening instruments for women (the Correctional Mental Health Screen–Female [CMHS-F]) and men (the Correctional Mental Health Screen–Male [CMHS-M]) that may provide an efficient, empirically-based approach to the critical first stage of detecting clinically significant mental health problems that otherwise are likely to be undetected among newly incarcerated adults.

## METHOD

### Participants

The study participants were 2,196 adults (1,526 men and 670 women aged 18 to 78 years [$M$ = 32 years, $SD$ = 9.5 years]) admitted into Connecticut jail within the previous 24 to 72 hours. Criminal charge data were obtained from Connecticut Department of Correction (DOC) records for the 85% of the sample ($n$ = 1,868) who consented to release this information to the study. Men were recruited in the four jails for male inmates in the state correction system. Women were recruited in the state's single correctional facility for women, both unsentenced and after sentencing by the courts. Participants' self-reported ethnicities were White, not Hispanic (41%), Black (including African American and Caribbean American; 38%), Latino or Hispanic (20%), Asian American (0.5%), and American Indian (0.2%). Educational levels ranged from 0 to 19 years of school ($M$ = 11.4 years, $SD$ = 1.7 years). Primary criminal charges included nonviolent crimes (58%), violent crimes (16%), and probation violations (26%). The demographic composition and criminal charges of the sample were consistent with characteristics of the overall jail population, except for ethnicity: Hispanics were slightly underrepresented (20% screened vs. 27% prevalence in Connecticut jails).

### Procedure

Screenings were conducted to provide a sample that included new admissions Monday through Friday afternoons and evenings, on a schedule dictated by each facility's logistics and policies. When RAs arrived at each facility, an intake list was provided by custody supervisors that included the age, ethnicity, and correctional status of each admission in the past 24 (Tuesday to Friday) or 72 (Monday) hours. Adults aged 18 years or older were (a) identified in the general population (i.e., not confined to restricted units because of acute dangerousness), (b) not "high-bond" security risks (because these individuals could not be interviewed without custody officers present), and (c) not already admitted to medical units (for acute treatment of wounds or injuries or acute substance intoxication or detoxification) or to mental health units (for severe acute psychosis, suicidality, mania, delirium, or documented histories of intensive psychiatric care).

The RAs randomly selected inmates from the eligible individuals on the intake lists, with one exception. After reviewing the first 1,000 screenings, we found that Whites were overrepresented and Hispanics were underrepresented. The sampling strategy was modified to oversample Hispanics and undersample Whites for the remaining 1,196 study screenings.

Custody officers escorted selected inmates to specially designated private interview rooms (usually in the medical units) to meet with the RAs. Following a brief description of the study purpose and participation requirements, the RAs provided willing inmates with written and oral descriptions of the informed-consent process. To ensure comprehension, each person was asked to describe, in his or her own words, the study's purpose, what he or she was being asked to do, that declining to participate would not affect incarceration or services, and that participation could be discontinued at any time. Participants were informed that they would not receive compensation of any kind. Inmates who chose not to participate were thanked by the

RAs and escorted to their residential units by custody officers. The consent process and form and all study procedures, personnel, and tests were reviewed and approved by an appropriately constituted institutional review board of the University of Connecticut Health Center and by the Connecticut DOC Research Advisory Committee.

Consenting persons were given a copy of the composite screen to read while RAs read each question and answer options orally and recorded the participants' oral answers. The order of presentation of the screen's subscales was randomly varied. RAs briefly redefined any words or phrases that were not understood by participants.

The screening RAs used a preset numerical sequence to randomly select one in five participants for invitation to participate in the follow-up interviews with different RAs. Neither the screening RAs nor the participants knew whether the participants would be invited to do the follow-up interviews until the screening interviews were completed. Informed consent was fully readministered at the start of the follow-up interviews. The order of measures in the follow-up interviews was invariant to enhance participants' retention and completed interviews: first *DSM-IV* Axis I disorders other than PTSD, then PTSD, and finally *DSM-IV* Axis II disorders.

There were two exceptions to the randomized 20% follow-up interview sampling procedure. If a participant declined to do the follow-up interview, the subsequent screening participant was invited. The first 1,050 screening interviews were assigned to follow-up interviews on a 10% basis because of an insufficient number of assessors, leading to an effective 14% sampling overall ($n$ = 302). There were no differences in age, education level, and criminal charges between interview participants and all jail admissions, but Whites were overrepresented (43% vs. 28% in the jail census) and Blacks and Hispanics (35% and 22% vs. 44% and 27%, respectively, in the jail census) were underrepresented. The disparity may in part be because although we were able to recruit a bilingual, Spanish-speaking screener, we were unable to do so for follow-up interviews.

Interview RAs who were blind to screening results conducted follow-up interviews within the next 5 days in private interview rooms in the same areas where the screening interviews took place. RAs read all interview questions aloud and provided visual aids to facilitate numerical responses. As in the screening, interview RAs briefly redefined any words or phrases not understood and probed only sufficiently to get a "yes" or "no" response to each interview item. The follow-up interview required 45 to 180 minutes to complete (median = 130 minutes).

The final sample did not differ on any demographic or arrest characteristics from the overall DOC population, except that White inmates (41% vs. 28% overall) were overrepresented and Black (37% vs. 44% overall) and Hispanic (21% vs. 27% overall) inmates were underrepresented. One in four potential participants (24%) were ineligible, including 99 (1%) who did not speak English, and inmates who were in court, high bond or high risk, bonded out, in the hospital, admitted to inpatient mental health, under 18 years of age, in restricted housing, under transfer to other units or facilities, or being held under the custody of immigration or for federal offenses. Ninety percent of the eligible invitees participated in the screening, and 100% of those invited to participate in the interviews did so. Those who declined gave the following reasons for refusal: in detoxification; preferred to be in the gym, doing recreation, or eating meals; feeling ill; and unable to speak English. Gender, age, and ethnicity were unrelated to refusal, except that more Black women (51% of refusals) than White women (36% of refusals) refused, $\chi^2(2) = 13.99$, $p < .003$.

The BA-level RAs who conducted screening interviews included three White women (range = 101 to 723 screenings of men and 61 to 432 screenings of women per interviewer), one Latina (97 screenings of men), and one White man (239 screenings of men). RAs were trained and supervised in the administration of the screening interview by the authors (JDF, VW), including extensive role-play practice with coaching and feedback and observing at least 10 screening interviews firsthand prior to conducting interviews alone. The BA-, MA-, and MD-level RAs who conducted follow-up interviews included four White women (range = 12 to 65 interviews with men per interviewer and 6 to 55 with women) and three White men (range = 8 to 22 interviews with men). RAs were trained and supervised by the first author (i.e., role-play practice with coaching and feedback, at least 3 interviews in situ prior to independent interviews). Screenings and interviews were reviewed weekly in supervision. No adverse reactions by participants were observed by the RAs or reported by custody or correctional mental health staff members during or after any screening or interview.

Interrater reliability was assessed for 124 screening interviews and 16 follow-up interviews conducted in a second phase of the study with secondary interviewers silently observing and independently recording numerical answers. Temporal stability was assessed by retests with 100 men and 100 women between 1 and 5 days after the initial screening.

## Composite Screening Measure

Five psychometrically robust questionnaires constituted the screening interview, four of which were used to yield

56 dichotomous scores representing a broad spectrum (to maximize content validity) of the primary criterion symptoms for *DSM-IV* Axis I mood disorders (bipolar, major depressive, and dysthymic disorders), psychotic disorders (schizophrenia), anxiety disorders (panic disorder, agoraphobia, PTSD, obsessive-compulsive disorder, generalized anxiety disorder, and social phobia), somatoform disorders (somatization disorder, hypochondriasis), eating disorders (anorexia, bulimia), and the major features of Clusters A, B, and C Axis II psychiatric disorders.

*Screening module from the SCID–Patient Version (SCID-P; First, Spitzer, Gibbon, & Williams, 1996)*. This 24-item screen is designed to rule out disorders rapidly (7- to 10-minute administration time) using lead criteria for each Axis I disorder. Items provide dichotomous scores for depression and bipolar disorder (2 items each); dysthymia (1 item); panic, agoraphobia, social phobia, and generalized anxiety (1 item each); obsessive-compulsive disorders (2 items); psychotic disorders (8 items); eating disorders (4 items); and somatization disorder (1 item).

*Primary Care PTSD Screen (Prins et al., 2003)*. This 4-item symptom self-report measure addressing PTSD has demonstrated reliability, validity, and diagnostic utility with adult primary care populations.

*Iowa Personality Disorders Screen (Langbehn et al., 1999)*. This 14-item measure is derived from the Structured Interview for *DSM* Personality Disorders for adults in clinical and community samples. Five- and 7-item combinations from the Iowa Personality Disorders Screen prospectively identified psychiatric outpatients and inpatients with personality disorders with sensitivity of 0.79 to 0.92 and specificity of 0.79 to 0.86. Items represent primary symptoms of Axis II Clusters A, B, and C.

*Referral Decision Scale (Teplin & Swartz, 1989)*. The RDS is a temporally reliable 14-item measure that has shown mixed evidence for validity and predictive utility (as described above) for identifying jail inmates with bipolar, major depression, and psychotic disorders.

## Structured Diagnostic Interview

The follow-up structured interview included the SCID-P (First et al., 1996) for Axis I and II disorders. The Clinician Administered PTSD Scale (Weathers, Ruscio, & Keane, 1999) provided a reliable and validated PTSD diagnostic measure, supplemented by questions adapted from the Stressful Life Experiences Scale (Goodman, Corcoran, Turner, Yuan, & Green, 1998) and the Potential Stressful Events Interview (Resnick, Falsetti, Kilpatrick, & Freedy, 1996) to assess PTSD Criterion A.

## Correctional Records Data

Research assistants trained by correctional facility administrators and supervised by the second author reviewed computerized offender databases to extract data on the basis of risk and needs assessments (on a scale ranging from 1 = *low risk/need* to 5 = *maximal risk/need*) rated by correctional counselors trained in their use by the Connecticut DOC Division of Population Management. Ratings were made according to guidelines from DOC administrative directives (Connecticut Department of Correction, 2005) that were designed to be consistent with standards set by the American Correctional Association. Risks included one hypothesized to be related to psychiatric illness (suicide) and others unrelated to psychiatric illness (potential violence, disruptive behavior, or escape). Needs included ones hypothesized to be related to psychiatric illness (mental health care and substance abuse treatment) and others unrelated to psychiatric illness (medical care, vocational and educational training, sexual offender treatment, and family and community resources).

## Statistical Analyses

We first conducted item reduction analyses separately for each gender, eliminating items with low (<10%) or high (>90%) base rates. Although low base-rate items might have utility in ruling in or out diagnoses or syndromes that also have low base rates, we chose not to include them in this study because they would be too infrequently endorsed in the gender-by-ethnicity subgroups of the interview sample to provide a reliable basis for estimating their predictive utility. Bivariate Pearson correlations were computed for all remaining pairs of items, and sets of items with $r > .50$ were reduced to a single item by random selection if their content was judged redundant (unless two items each had high accuracy for at least two different sets of diagnostic criteria, in which case both were retained; see below). Measures of sampling adequacy (MSAs; Cerny & Kaiser, 1977) were computed to further reduce the item sets; items with low MSAs are largely uncorrelated with the full set. The items in the resultant sets of 40 variables for men (MSA $\geq 0.90$) and 38 variables for women (MSA $\geq 0.85$) all exceeded the recommended cutoff of 0.80 for sampling adequacy (Cerny & Kaiser, 1977).

The sensitivity, specificity, PPP, and NPP (Kessel & Zimmerman, 1993) of each remaining item were tested for predicting the presence of any psychiatric diagnosis. Current psychiatric disorders were less common among men (56% prevalence) than women (68% prevalence), and both sensitivity and specificity were lower for identifying

**TABLE 1**
**Brief Mental Health Screens for Newly Incarcerated Women (CMHS-F) and Men (CMHS-M)**

| *Item* | | *PPP* |
|---|---|---|
| CMHS-F | | |
| RDSq10: | Ever hospitalized for psychiatric reasons | 1.00 |
| RDSq14: | Felt useless, sinful or guilty few weeks in a row | 0.92 |
| SCIDq3: | Feel depressed most of the day, every day | 0.88 |
| IPDSq1b: | Mood changes frequently during day (emotional rollercoaster) | 0.87 |
| PTSDq3: | Repeated thoughts or feelings about stressful past experience | 0.85 |
| IPDSq4: | People will take advantage if they get to know you | 0.84 |
| IPDSq11: | Annoyed when friends and family complain | 0.84 |
| PTSDq4: | Tried to avoid reminders or not think about stressful past event | 0.83 |
| CMHS-M | | |
| RDSq10: | Ever hospitalized for psychiatric reasons | 0.90 |
| IPDSq1b: | Mood changes frequently during day (emotional rollercoaster) | 0.85 |
| IPDSq3b: | Trouble at work because excited but do not follow through | 0.82 |
| SCIDq3: | Feel depressed most of the day, every day | 0.80 |
| SCIDq17: | Worries that can't be gotten rid of | 0.80 |
| PTSDq3: | Repeated thoughts or feelings about stressful past experience | 0.78 |
| PTSDq5: | No feelings or distant and cut off from people or surroundings | 0.75 |
| PTSDq6: | Constantly on guard/watchful when not necessary or easily startled | 0.75 |
| SCIDq5: | Irritable most of the time | 0.73 |
| PTSDq4: | Tried to avoid reminders or not think about stressful past event | 0.72 |
| IPDSq10: | Tend to hold grudges or give silent treatment for days at a time | 0.72 |
| IPDSq11: | Annoyed when friends and family complain | 0.67 |

NOTE: CMHS-F = Correctional Mental Health Screen–Female; CMHS-M = Correctional Mental Health Screen–Male; PPP = positive predictive power for identifying psychiatric diagnosis; RDSq = Referral Decision Scale items; PTSDq = Primary Care PTSD Screen items; SCIDq = Structured Clinical Interview for *DSM-IV*–Patient Version items; IPDSq = Iowa Personality Disorders Screen items.

men with psychiatric disorders than for identifying women. Items with PPP ≥ 0.80 for women and PPP ≥ 0.67 for men were retained in two provisional brief screens, one for women (8 items; CMHS-F) and one for men (12 items; CMHS-M) (Table 1).

Interrater and retest reliability were assessed for each CMHS-F and CMHS-M item by calculating κ and for the total CMHS-F and CMHS-M with intraclass correlations. Cronbach's α was used to assess the internal consistency of the CMHS-F and CMHS-M.

Discriminant function analyses tested the predictive utility of the CMHS-F and CMHS-M in relation to five categories of psychiatric disorders: (a) any current Axis I disorder, (b) any psychiatric disorder (current Axis I or Axis II), (c) a psychiatric disorder except ASPD, (d) a current affective disorder, and (e) a current anxiety disorder. Most Axis I and II disorders were uncommon in this nonclinical population, and distinctions among disorders have less relevance in the initial screening phase of correctional mental health care than whether a disorder of a given class is present (Steadman et al., 2005). The selected categories of disorders share common features within the *DSM-IV*, were prevalent (>10% for men, >33% for women) in this sample, and are associated with emotional or behavioral instability that is relevant to correctional management (including risk for harming oneself or others, as well as problems adhering to the firmly controlled activity schedule and disciplinary standards in the first 14 days of incarceration).

ROC analyses were used next to examine predictive utility (accuracy, sensitivity, specificity, PPP, NPP, and area under the curve [AUC]) for optimal cut points for the CMHS-F and CMHS-M using the closest-to-(0,1) criterion and the Youden index (*J*) (Perkins & Schisterman, 2006). The closest-to-(0,1) criterion identifies the cut point on the ROC curve that is the closest to the ideal of perfect specificity (0 on the *x* axis) and sensitivity (1 on the *y* axis). The Youden index uses the opposite approach by selecting the cut point on the ROC curve that is furthest from a purely chance association of the predictor and criterion (i.e., the greatest distance from the diagonal line on the ROC graph that reflects only chance predictive accuracy). If the optimal classification cut points for any subsample consistently led to NPP < 0.70 (i.e., missing >30% of the true cases), supplemental analyses were conducted using alternative cut points that resulted in maximal NPP.

The incremental utility of the CMHS was tested by comparing its performance with that of the RDS and BJMHS (a) using the full range of scale scores and ROC-derived

optimal cut points to predict the five psychiatric disorder categories and (b) using the full range of scores on each screener and standardized residuals from linear regression analyses to represent the unique contribution of each screener to the prediction of an undetected psychiatric disorder (Edens, Skeem, & Douglas, 2006). The latter test was done with all five of the psychiatric disorder categories (results available on request from the first author) but are reported only for the category of psychiatric disorders other than ASPD to simplify the results and focus on disorders that are more treatable yet more often overlooked than ASPD.

Finally, several tests of the validity of the CMHS were conducted. Criterion validity was tested with independent-samples *t* tests comparing mean CMHS scores for respondents who did and did not meet diagnostic criteria. CMHS-F and CMHS-M convergent and discriminant validity were tested with bivariate correlations between their total scores with correctional records variables that represented, respectively, mental health–relevant and non–mental health constructs.

## RESULTS

### Reliability of the CMHS-F and CMHS-M

An 8-item screen for women (CMHS-F) and a 12-item screen for men (CMHS-M) were created (see Table 1 for a description of items) on the basis of the item reduction procedures described above. Internal consistency (Cronbach's $\alpha$) was .76 for the CMHS-F and .78 for the CMHS-M. For the CMHS-F, retest $\kappa$ values (for each item) were .41 to .91 (median = .66). The only $\kappa$ value less than .55 was for the item assessing PTSD avoidance symptoms: 8% to 15% of respondents reversed their first responses on the second screening. The retest Pearson *r* value for the total CMHS-F score was .82, and the more conservative Kendall's $\tau$-c value (a test for an excess of concordant over discordant cases, adjusted for table size) was .70. For the CMHS-M, retest $\kappa$ values (for each item) were .47 to .83 (median = .62). The two items for which $\kappa$ was below .50 assessed PTSD avoidance symptoms and generalized anxiety; for both, 12% to 14% of respondents reversed their responses at the second testing. The retest Pearson *r* value for the total CMHS-M score was .84, and Kendall's $\tau$-c value was .73. Interrater agreement for each item in the CMHS-F and CMHS-M and for the interview diagnosis categories was 100%.

### Predictive Utility of the CMHS-F

Discriminant function analyses for the CMHS-F were statistically significant for identifying (a) any current Axis I diagnosis, $\Lambda = .727$, $\chi^2(8, N = 100) = 30.02$, $p < .001$ (73.0% correctly classified); (b) any current Axis I or Axis II psychiatric diagnosis, $\Lambda = .746$, $\chi^2(8, N = 100) = 27.55$, $p = .001$ (70.0% correctly classified); (c) any psychiatric diagnosis except ASPD, $\Lambda = .661$, $\chi^2(8, N = 100) = 38.88$, $p < .001$ (77.0% correctly classified); (d) any current affective disorder, $\Lambda = .717$, $\chi^2(8, N = 99) = 30.90$, $p < .001$ (73.7% correctly classified); and (e) any current anxiety disorder, $\Lambda = .751$, $\chi^2(8, N = 97) = 26.05$, $p = .001$ (73.2% correctly classified).

ROC analyses (Figure 1) for all women, and separately for subsamples of White and Black women (too few Hispanic women were tested to provide reliable results), yielded virtually identical cut points for optimal classification with the closest-to-(0,1) criterion and the *J* index (Table 2). Using cut points based on the *J* index, the CMHS-F accounted for statistically significant and clinically meaningful AUCs (median = 0.75; see Table 2). In two cases, the AUCs were less than 0.70: For Black women, AUC = 0.64 for affective disorders, with perfect sensitivity and NPP but weak specificity and PPP, and AUC = 0.69 for any Axis I or II disorder other than ASPD, with less than 0.70 sensitivity and NPP and PPP but poorer (0.58) specificity. Using the optimal cut point, specificity was particularly robust for White women across disorder categories (range = 0.83 to 1.00, median = 0.92), while for Black women, sensitivity was strongest (range = 0.75 to 1.00, median = 1.00). Consistent with these findings, the optimal classification cut points were three points lower (2 to 3) for Black women than for the full sample and for White women (5 to 6).

Given the less than optimal ability of the CMHS-F to identify true cases among White women (i.e., NPP = 0.49 to 0.84 with the *J* index cut points), supplemental analyses identified alternative cut points to maximize NPP for the full female sample and for White women (see Table 2; maximal NPP cut points are not shown for Black women because for them, the optimal classification cut points yielded maximal NPP levels). With one exception (identifying anxiety disorder cases in the White female subsample), lower CMHS-F cut points (1 to 2), comparable with the optimal cut points for Black women, yielded maximal NPP (0.83 to 1.00) for identifying all diagnostic categories except when ASPD was included. Use of the maximal-NPP CMHS-F cut points would lead to up to 31% additional false positive (median = 21%), compared with using the optimal classification cut point.

In comparison with the results with other diagnostic categories, the CMHS-F did not accurately identify ASPD cases among women. The maximal NPP for identifying a current Axis I or Axis II disorder, including ASPD, was low for the full sample and the White and Black subsamples (0.50 to 0.67).

**TABLE 2**
**Predictive Utility of Screening Measures for Female Inmates With Cut Points to Optimize Classification (Youden index) or to Maximize the Identification of Undetected Cases (Maximal NPP)**

| *Cut Point* | *AUC* | *Sensitivity* | *Specificity* | *PPP* | *NPP* |
|---|---|---|---|---|---|
| **Full sample of female inmates** | | | | | |
| Affective disorder (current prevalence: 33.3%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 5 | 0.71** | 0.73 | 0.70 | 0.55 | 0.84 |
| RDS 6 | 0.70** | 0.61 | 0.79 | 0.59 | 0.80 |
| BJMHS 2 | 0.69** | 0.94 | 0.44 | 0.48 | 0.93 |
| Maximal NPP cut point | | | | | |
| CMHS 2 | 0.59 | 1.00 | 0.20 | 0.38 | 1.00 |
| RDS 2 | 0.65* | 0.94 | 0.36 | 0.43 | 0.92 |
| BJMHS 2 | 0.69** | 0.94 | 0.44 | 0.48 | 0.93 |
| Anxiety disorder (current prevalence: 34.0%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 6 | 0.70** | 0.67 | 0.73 | 0.56 | 0.81 |
| RDS 6 | 0.69** | 0.61 | 0.77 | 0.57 | 0.79 |
| BJMHS 5 | 0.69** | 0.52 | 0.87 | 0.67 | 0.78 |
| Maximal NPP cut point | | | | | |
| CMHS 2 | 0.57 | 0.97 | 0.19 | 0.38 | 0.92 |
| RDS 1 | 0.57 | 0.97 | 0.17 | 0.38 | 0.92 |
| BJMHS 1 | 0.57 | 0.94 | 0.20 | 0.37 | 0.86 |
| Any Axis I disorder (current prevalence: 43.0%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 6 | 0.71*** | 0.63 | 0.79 | 0.69 | 0.74 |
| RDS 2 | 0.67** | 0.93 | 0.40 | 0.54 | 0.89 |
| BJMHS 2 | 0.68** | 0.90 | 0.46 | 0.56 | 0.86 |
| Maximal NPP cut point | | | | | |
| CMHS 2 | 0.58 | 0.98 | 0.21 | 0.48 | 0.92 |
| RDS 1 | 0.59 | 0.98 | 0.21 | 0.48 | 0.92 |
| BJMHS 2 | 0.68** | 0.90 | 0.46 | 0.56 | 0.86 |
| Any current Axis I or Axis II disorder (prevalence: 68.0%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 6 | 0.74*** | 0.54 | 0.94 | 0.95 | 0.49 |
| RDS 6 | 0.69** | 0.47 | 0.91 | 0.91 | 0.45 |
| BJMHS 3 | 0.68** | 0.67 | 0.69 | 0.83 | 0.48 |
| Maximal NPP cut point | | | | | |
| CMHS 1 | 0.52 | 0.97 | 0.13 | 0.70 | 0.67 |
| RDS 2 | 0.64* | 0.84 | 0.47 | 0.77 | 0.58 |
| BJMHS 2 | 0.68** | 0.80 | 0.55 | 0.80 | 0.55 |
| Any current Axis I or Axis II disorder other than ASPD (prevalence: 60.0%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 5 | 0.75*** | 0.65 | 0.85 | 0.87 | 0.62 |
| RDS 2 | 0.70** | 0.90 | 0.50 | 0.73 | 0.77 |
| BJMHS 2 | 0.74*** | 0.88 | 0.60 | 0.77 | 0.76 |
| Maximal NPP cut point | | | | | |
| CMHS 1 | 0.53 | 0.98 | 0.13 | 0.63 | 0.83 |
| RDS 1 | 0.62 | 0.97 | 0.28 | 0.67 | 0.85 |
| BJMHS 1 | 0.62* | 0.95 | 0.30 | 0.68 | 0.79 |
| **White female inmate subsample** | | | | | |
| Affective disorder (current prevalence: 39.7%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 6 | 0.74** | 0.65 | 0.83 | 0.71 | 0.78 |
| RDS 8 | 0.69* | 0.44 | 0.94 | 0.83 | 0.72 |
| BJMHS 2 | 0.70* | 1.00 | 0.40 | 0.56 | 1.00 |
| Maximal NPP cut point | | | | | |
| CMHS 2 | 0.57 | 1.00 | 0.17 | 0.44 | 1.00 |
| RDS 2 | 0.65 | 1.00 | 0.31 | 0.49 | 1.00 |
| BJMHS 2 | 0.70* | 1.00 | 0.40 | 0.56 | 1.00 |

**TABLE 2 (continued)**

| *Cut Point* | *AUC* | *Sensitivity* | *Specificity* | *PPP* | *NPP* |
|---|---|---|---|---|---|
| Anxiety disorder (current prevalence: 32.8%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 6 | 0.77** | 0.74 | 0.80 | 0.64 | 0.86 |
| RDS 6 | 0.74** | 0.68 | 0.80 | 0.62 | 0.84 |
| BJMHS 5 | 0.75** | 0.65 | 0.86 | 0.69 | 0.84 |
| Maximal NPP cut point | | | | | |
| Optimal cut point | | | | | |
| CMHS 6 | 0.77** | 0.74 | 0.80 | 0.64 | 0.86 |
| RDS 1 | 0.57 | 1.00 | 0.15 | 0.37 | 1.00 |
| BJMHS 3 | 0.66 | 0.82 | 0.50 | 0.44 | 0.86 |
| Any Axis I disorder (current prevalence: 47.5%) | | | | | |
| CMHS 6 | 0.76** | 0.64 | 0.87 | 0.82 | 0.73 |
| RDS 6 | 0.71** | 0.57 | 0.84 | 0.76 | 0.68 |
| BJMHS 5 | 0.66* | 0.46 | 0.86 | 0.75 | 0.63 |
| Maximal NPP cut point | | | | | |
| CMHS 2 | 0.53 | 0.96 | 0.16 | 0.51 | 0.83 |
| RDS 1 | 0.59 | 1.00 | 0.19 | 0.53 | 1.00 |
| BJMHS 1 | 0.59 | 0.96 | 0.21 | 0.53 | 0.86 |
| Any current Axis I or Axis II psychiatric disorder (prevalence: 67.8%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 6 | 0.78** | 0.55 | 1.00 | 1.00 | 0.51 |
| RDS 6 | 0.69* | 0.48 | 0.90 | 0.91 | 0.45 |
| BJMHS 5 | 0.67 | 0.40 | 0.94 | 0.94 | 0.40 |
| Maximal NPP cut point | | | | | |
| CMHS 2 | 0.54 | 0.95 | 0.21 | 0.72 | 0.67 |
| RDS 2 | 0.56 | 0.88 | 0.32 | 0.73 | 0.55 |
| BJMHS 2 | 0.61 | 0.84 | 0.38 | 0.76 | 0.50 |
| Any current Axis I or Axis II disorder other than ASPD (prevalence: 62.7%) | | | | | |
| Optimal cut point | | | | | |
| CMHS 5 | 0.81*** | 0.70 | 0.91 | 0.93 | 0.65 |
| RDS 6 | 0.71** | 0.51 | 0.91 | 0.91 | 0.53 |
| BJMHS 2 | 0.69* | 0.91 | 0.47 | 0.76 | 0.75 |
| Maximal NPP cut point | | | | | |
| CMHS 2 | 0.57 | 0.97 | 0.23 | 0.68 | 0.83 |
| RDS 1 | 0.63 | 1.00 | 0.27 | 0.70 | 1.00 |
| BJMHS 1 | 0.64 | 0.97 | 0.32 | 0.72 | 0.86 |
| **Black female inmate subsample** | | | | | |
| Affective disorder (current prevalence: 22.2%) | | | | | |
| CMHS 2 | 0.64 | 1.00 | 0.29 | 0.29 | 1.00 |
| RDS 5 | 0.69 | 0.67 | 0.71 | 0.40 | 0.88 |
| BJMHS 4 | 0.71 | 0.67 | 0.76 | 0.44 | 0.89 |
| Anxiety disorder (current prevalence: 36.0%) | | | | | |
| CMHS 3 | 0.78* | 1.00 | 0.56 | 0.56 | 1.00 |
| RDS 7 | 0.78* | 0.56 | 1.00 | 1.00 | 0.80 |
| BJMHS 2 | 0.70 | 0.78 | 0.63 | 0.54 | 0.83 |
| Any current Axis I disorder (prevalence: 37.0%) | | | | | |
| CMHS 3 | 0.79* | 1.00 | 0.59 | 0.59 | 1.00 |
| RDS 2 | 0.77* | 0.90 | 0.65 | 0.60 | 0.92 |
| BJMHS 2 | 0.75* | 0.80 | 0.71 | 0.62 | 0.86 |
| Any current Axis I or Axis II psychiatric disorder (prevalence: 74.1%) | | | | | |
| CMHS 3 | 0.73 | 0.75 | 0.71 | 0.88 | 0.50 |
| RDS 2 | 0.88** | 0.75 | 1.00 | 1.00 | 0.58 |
| BJMHS 2 | 0.83* | 0.65 | 1.00 | 1.00 | 0.50 |
| Any current Axis I or Axis II disorder other than ASPD (prevalence: 55.6%) | | | | | |
| CMHS 3 | 0.69 | 0.80 | 0.58 | 0.71 | 0.70 |
| RDS 3 | 0.78* | 0.73 | 0.83 | 0.85 | 0.71 |
| BJMHS 2 | 0.78* | 0.73 | 0.83 | 0.85 | 0.71 |

NOTE: $N = 101$ women. NPP = negative predictive power; AUC = area under the curve; PPP = positive predictive power; CMHS = Correctional Mental Health Screen; RDS = Referral Decision Scale; BJMHS = Brief Justice Mental Health Screen; ASPD = antisocial personality disorder. Numbers following CMHS, RDS, and BJMHS refer to the scores designated as the cut points. Predictive utility statistics for other cut points may be obtained from the first author by request. All values for Black female subsample are based on optimal cut-points which did not differ from maximal NPP cut-points except in two instances (available from first author).
*$p < .05$. **$p < .01$. ***$p < .001$.

**FIGURE 1**
**Receiver-Operating Characteristic (ROC) Analysis: Areas Under the Curves for Identifying Female Jail Inmates With Axis I or II Psychiatric Disorders (Excluding Antisocial Personality Disorder), With the Correctional Mental Health Scale–Female (CMHS-F; Solid Line), the Brief Jail Mental Health Screen (BJMHS; Dotted Line), and the Referral Decision Scale (RDS; Dashed Line)**

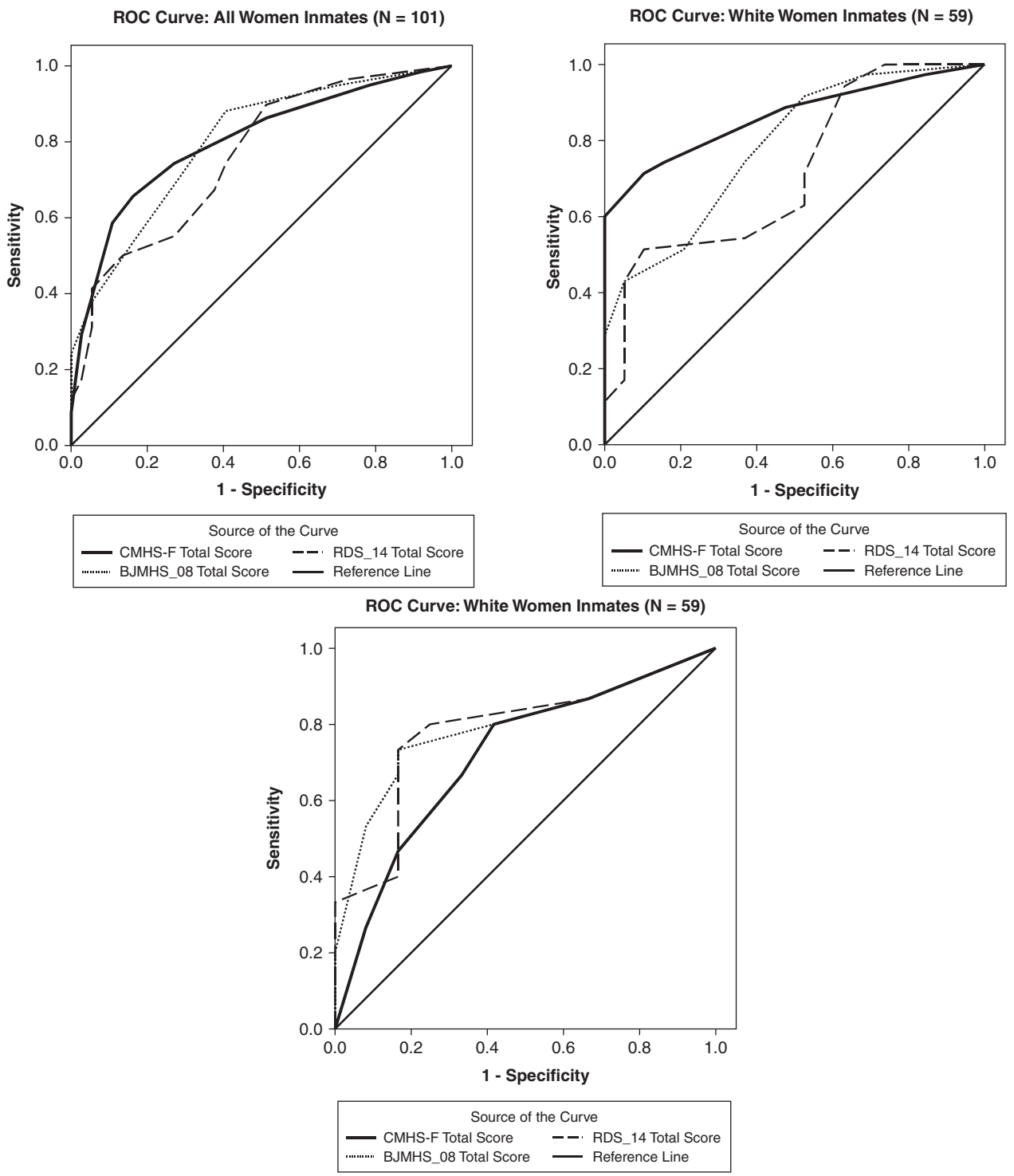


**TABLE 3**
**Predictive Utility of the CMHS, RDS, and BJMHS Across the Full Range of Scores for Identifying Undetected Axis I or II Psychiatric Disorder (Excluding Antisocial Personality Disorder)**

| *Screening Measure* | *AUC* | SE | *95% Confidence Interval* |
|---|---|---|---|
| Full sample: women | | | |
| CMHS-F | 0.80 | 0.05 | 0.71 to 0.89 |
| RDS | 0.76 | 0.05 | 0.66 to 0.86 |
| BJMHS | 0.80 | 0.05 | 0.71 to 0.89 |
| White women | | | |
| CMHS-F | 0.86 | 0.05 | 0.77 to 0.96 |
| RDS | 0.72 | 0.07 | 0.57 to 0.86 |
| BJMHS | 0.78 | 0.06 | 0.66 to 0.91 |
| Black women | | | |
| CMHS-F | 0.72 | 0.10 | 0.52 to 0.92 |
| RDS | 0.78 | 0.09 | 0.60 to 0.96 |
| BJMHS | 0.79 | 0.08 | 0.62 to 0.97 |
| Full sample: men | | | |
| CMHS-M | 0.79 | 0.04 | 0.72 to 0.86 |
| RDS | 0.69 | 0.04 | 0.61 to 0.76 |
| BJMHS | 0.70 | 0.04 | 0.62 to 0.77 |
| White men | | | |
| CMHS-M | 0.84 | 0.05 | 0.75 to 0.94 |
| RDS | 0.70 | 0.06 | 0.60 to 0.81 |
| BJMHS | 0.73 | 0.05 | 0.63 to 0.83 |
| Black men | | | |
| CMHS-M | 0.77 | 0.06 | 0.72 to 0.33 |
| RDS | 0.67 | 0.07 | 0.53 to 0.79 |
| BJMHS | 0.63 | 0.07 | 0.50 to 0.77 |

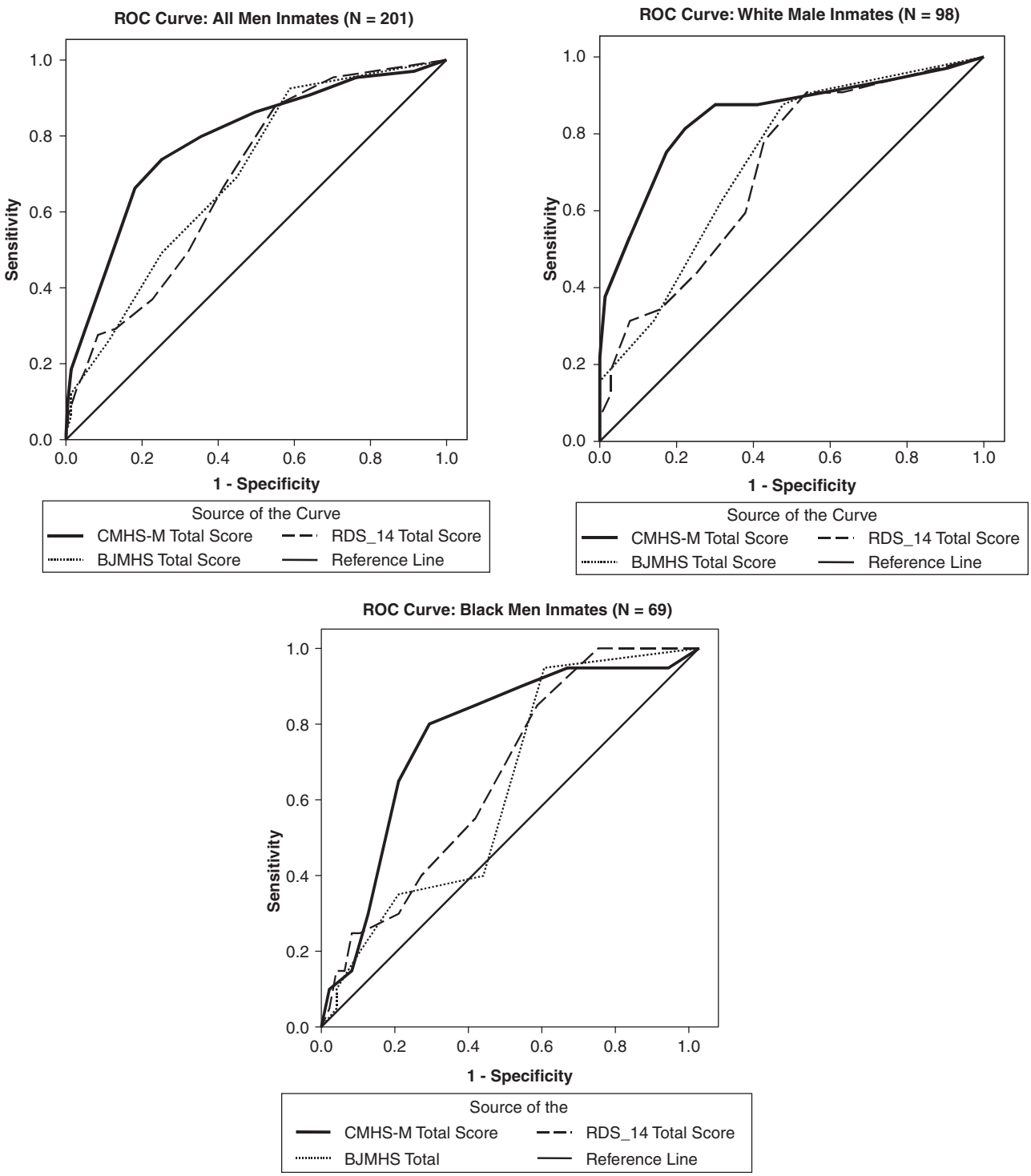
NOTE: N = 101 women, 201 men. CMHS = Correctional Mental Health Screen; RDS = Referral Decision Scale; BJMHS = Brief Justice Mental Health Screen; AUC = area under the curve. Predictive utility statistics for the identification of the four other categories of psychiatric disorders (i.e., any anxiety disorder, any affective disorder, any current Axis I disorder, any current Axis I or Axis II disorder) may be obtained from the first author by request.

We next conducted ROC analyses comparing the CMHS-F with items from the composite screen representing the 14-item RDS (Teplin & Swartz, 1989) and the 8-item BJMHS (Steadman et al., 2005) with the full sample and White and Black female subsamples (Figure 1). Optimal cut points for each screener were identified, and the AUCs, sensitivity, specificity, and PPP and NPP for each are displayed in Table 2. Full-sample AUCs based on the optimal classification cut points were higher for the CMHS-F (median = 0.73) than for the RDS (median = 0.69) and the BJMHS (median = 0.69), for all five psychiatric disorder categories. Using the full range of scores, the CMHS-F and BJMHS had identical AUCs for the prediction of any Axis I or II disorder, excluding ASPD, in the full female sample, higher than that for the RDS (Table 3). In the White female subsample, AUCs based on optimal classification cut points with the CMHS-F (median = 0.77) also exceeded those of the RDS (median = 0.71) and the BJMHS (median = 0.69) for the five psychiatric disorder categories. The CMHS-F also had higher AUCs than the RDS or BJMHS among White women when the full range of scores was used to predict any Axis I or II disorder, excluding ASPD (Table 3). The NPP levels (and the increase in false positives) obtained when using maximal-NPP cut points with the CMHS-F were comparable with the NPP levels obtained when maximal-NPP cut points were calculated for the BJMHS and RDS items (Table 2).

With Black women, the AUCs using optimal cut points were higher for the CMHS-F than the RDS or BJMHS for current anxiety disorders or any current Axis I disorder but higher for the RDS and BJMHS than the CMHS-F for current affective disorders and an Axis I or II disorder with or without ASPD (Table 2). Using the full range of scores to predict any Axis I or II disorder, excluding ASPD, the CMHS-F had a lower AUC than the BJMHS or RDS with Black women (Table 3).

As a test of CMHS-F incremental validity, ROC analyses calculated for the full sample of women with standardized residuals derived from linear regression analyses to predict the presence of any psychiatric disorder (other than ASPD) showed the CMHS-F's unique contribution to be equal to that of the BJMHS (AUC = 0.62 for the residual reflecting the unique contribution of either the CMHS-F or the BJMHS) and significantly greater (AUC = 0.67, 95% confidence interval [CI] = 0.56 to 0.78) than that for the RDS (AUC = 0.54, 95% CI = 0.42 to 0.65, $p < .05$; see Figure 2, top left). With the White female subsample, the CMHS-F's unique contribution was greater ($p < .05$; see Figure 2, top right) than that of the BJMHS (AUC = 0.70, 95% CI = 0.55 to 0.84 vs. AUC = 0.54, 95% CI = 0.38 to 0.70) or the RDS (AUC = 0.77, 95% CI = 0.66 to 0.90 vs. AUC = 0.41, 95% CI = 0.26 to 0.57). With the Black female subsample, the CMHS-F's unique contribution was lower than that of the BJMHS (AUC = 0.56, 95% CI = 0.34 to 0.78 vs. AUC = 0.71, 95% CI = 0.52 to 0.91) or the RDS (AUC = 0.55, 95% CI = 0.33 to 0.77 vs. AUC = 0.66, 95% CI = 0.44 to 0.87), but the differences were not statistically significant ($p > .10$). As illustrated in Figure 2, bottom the CMHS-F's accuracy with Black women was comparable with that of the RDS or the BJMHS at midlevel scores (specificity and sensitivity ~ 0.55 to 0.65) but poorer than that of the RDS or the BJMHS at the low or high end of the score levels. Similar findings were obtained in the full sample and the subsamples in parallel analyses conducted to predict an (a) affective disorder, (b) anxiety disorder, (c) Axis I disorder, and (d) Axis I or II disorder (results available from the first author).

**FIGURE 2**
**Receiver-Operating Characteristic (ROC) Analysis: Areas Under the Curves for Identifying Male Jail Inmates With Axis I or II Psychiatric Disorders (Excluding Antisocial Personality Disorder) With the Correctional Mental Health Scale–Male (CMHS-M; Solid Line), the Brief Jail Mental Health Screen (BJMHS; Dotted Line), and the Referral Decision Scale (RDS; Dashed Line)**




NOTE: ROC = receiver-operating characteristic.

## Predictive Utility of the CMHS-M

Discriminant function analyses for the CMHS-M were statistically significant for identifying (a) any current Axis I diagnosis, $\Lambda = .791$, $\chi^2(12, N = 200) = 44.90$, $p < .001$ (74.0% correctly classified); (b) any current Axis I or Axis II psychiatric diagnosis, $\Lambda = .792$, $\chi^2(12, N = 198) = 44.24$, $p < .001$ (72.7% correctly classified); (c) any current Axis I psychiatric diagnosis or Axis II diagnosis except ASPD, $\Lambda = .668$, $\chi^2(12, N = 198) = 71.17$, $p < .001$ (77.87% correctly classified); (d) any current affective disorder, $\Lambda = .735$, $\chi^2(12, N = 198) = 58.45$, $p < .001$ (81.37% correctly classified); and (e) any current anxiety disorder, $\Lambda = .865$, $\chi^2(12, N = 186) = 25.72$, $p < .05$ (74.2% correctly classified).

ROC analyses (Table 4, Figure 3) for all men and separately for the White and Black male subsamples yielded optimal classification cut points for each diagnostic outcome that were comparable on the basis of either the closest-to-(0,1) criterion or the Youden index. All CMHS-M AUCs were statistically significant (median = 0.74) and clinically significant (≥.70), except for identifying an Axis I or II disorder (total sample and both subsamples) or an anxiety disorder (total sample and White subsample). With optimal classification cut points, sensitivity and NPP (median = 0.75 and 0.85, respectively) were stronger than specificity and PPP (median = 0.70 and 0.62, respectively). Therefore, no supplementary analyses were run with the data for men to identify alternate cut points to maximize NPP. Optimal cut points were identical or within 1 point for the total sample and both subsamples, except for identifying an anxiety disorder (White male subsample cut point = 7, total male sample and Black subsample cut point = 4) and an Axis I or II disorder excluding ASPD (White male subsample cut point = 7, total male sample and Black subsample cut point = 5).

The full male sample AUCs using optimal cut points accounted for by the CMHS-M (median = 0.70) exceeded those for the RDS (median = 0.66) and BJMHS (median = 0.67) for all five composite psychiatric disorder categories, except that the BJMHS's AUC (0.67) for identifying an Axis I or II disorder, excluding ASPD, slightly exceeded the CMHS-M's AUC (0.66). Optimal cut point AUCs accounted for by the CMHS-M with White men (median = 0.77) exceeded those for the RDS (median = 0.71) and the BJMHS (median = 0.69). Optimal cut point AUCs for the CMHS-M with Black men (median = 0.74) also exceeded those of the RDS (median = 0.64) and BJMHS (median = 0.66). When the full range of scores was used to identify an Axis I or II disorder, excluding ASPD, the CMHS-M's AUCs exceeded those of the BJMHS or the RDS by 10% to 13% (Table 3).

As a test of CMHS-M incremental validity (calculated with standardized residuals derived from linear regression analyses), ROC analyses predicting the presence of any psychiatric disorder (other than ASPD) showed the CMHS-M's unique contribution (AUCs = 0.73 to 0.78, 95% CIs = 0.66 to 0.90 vs. the BJMHS; AUCs = 0.73 to 0.80, 95% CIs = 0.60 to 0.87 vs. the RDS) to be greater ($p < .05$) in the full sample of men and in the White and Black male subsamples than that of the BJMHS (AUCs = 0.43 to 0.53, 95% CIs = 0.27 to 0.66) or RDS (AUCs = 0.43 to 0.46, 95% CIs = 0.26 to 0.60). As illustrated in Figure 4, the CMHS-M was more accurate for the identification of both cases and noncases with midrange cut points than the BJMHS (with similar results for the RDS; figure available from the first author). Similar incremental validity findings for the CMHS-M were obtained in parallel analyses predicting the presence of a current (a) affective disorder, (b) anxiety disorder, (c) Axis I disorder, and (d) Axis I or II disorder (results available from the first author).

## Convergent, Discriminant, and Criterion Validity of the CMHS-F and CMHS-M

Table 5 displays bivariate Pearson correlations between the CMHS-F and CMHS-M total scores with correctional records data that served as indices of (a) mental health–relevant (i.e., convergent validity) and (b) non–mental health (i.e., discriminant validity) variables. Despite the fact that all participants had not been identified as having immediate mental health needs, there was variability in the mental health needs ratings (ranging from no needs to possible nonacute needs), and both the CMHS-F and CMHS-M total scores were correlated with mental health needs ratings. The CMHS-F score also was associated an inmate being ordered to have a mental health evaluation within the next 14 days after incarceration intake, but not with a correctional risk score for suicidality (data not available for men). The correlation of the CMHS-F score with correctional mental health needs ratings was statistically significantly stronger than the correlation of the CMHS-M score with the mental health needs rating (Fischer's exact statistic for $z$-transformed $r = 1.88$, $p < .05$).

With regard to discriminant validity, CMHS-F and CMHS-M scores were uncorrelated with correctional variables reflecting risk for violence, sex offenses, or escape; the severity of medical illness; vocational or educational needs; and history of substance abuse treatment, with one exception (Table 5). The exception was that the CMHS-F score weakly (<5% shared variance), but statistically

**TABLE 4**
**Predictive Utility of Screening Measures With Male Inmates With Cut Points to Optimize Classification (Youden Index)**

| *Cut Point* | *AUC* | *Sensitivity* | *Specificity* | *PPP* | *NPP* |
|---|---|---|---|---|---|
| **Full male inmate sample** | | | | | |
| Affective disorder (current prevalence: 12.1%) | | | | | |
| CMHS 7 | 0.78*** | 0.83 | 0.73 | 0.30 | 0.97 |
| RDS 3 | 0.69** | 0.83 | 0.55 | 0.20 | 0.96 |
| BJMHS 2 | 0.72*** | 0.92 | 0.53 | 0.21 | 0.98 |
| Anxiety disorder (current prevalence: 12.3%) | | | | | |
| CMHS 4 | 0.64* | 0.87 | 0.42 | 0.17 | 0.96 |
| RDS 2 | 0.62 | 0.87 | 0.37 | 0.16 | 0.95 |
| BJMHS 5 | 0.62 | 0.26 | 0.98 | 0.67 | 0.90 |
| Any Axis I disorder (current prevalence: 16.9%) | | | | | |
| CMHS 7 | 0.70*** | 0.68 | 0.72 | 0.33 | 0.92 |
| RDS 2 | 0.67** | 0.94 | 0.40 | 0.24 | 0.97 |
| BJMHS 2 | 0.68** | 0.82 | 0.53 | 0.26 | 0.94 |
| Any current Axis I or Axis II psychiatric disorder (prevalence: 56.3%) | | | | | |
| CMHS 5 | 0.66*** | 0.65 | 0.68 | 0.72 | 0.60 |
| RDS 2 | 0.63** | 0.77 | 0.48 | 0.66 | 0.62 |
| BJMHS 1 | 0.67*** | 0.85 | 0.49 | 0.68 | 0.71 |
| Any current Axis I or Axis II disorder other than ASPD (current prevalence: 33.7%) | | | | | |
| CMHS 6 | 0.75*** | 0.74 | 0.75 | 0.60 | 0.85 |
| RDS 2 | 0.66*** | 0.87 | 0.45 | 0.44 | 0.87 |
| BJMHS 1 | 0.67*** | 0.92 | 0.41 | 0.44 | 0.92 |
| **White male inmate subsample** | | | | | |
| Affective disorder (current prevalence: 17.7%) | | | | | |
| CMHS 7 | 0.86*** | 0.94 | 0.77 | 0.47 | 0.98 |
| RDS 3 | 0.74** | 0.94 | 0.54 | 0.31 | 0.98 |
| BJMHS 2 | 0.74** | 1.00 | 0.48 | 0.30 | 1.00 |
| Anxiety disorder (current prevalence: 10.9%) | | | | | |
| CMHS 7 | 0.69* | 0.70 | 0.68 | 0.21 | 0.95 |
| RDS 9 | 0.69 | 0.40 | 0.98 | 0.67 | 0.93 |
| BJMHS 5 | 0.69* | 0.40 | 0.99 | 0.80 | 0.93 |
| Any Axis I disorder (current prevalence: 20.4%) | | | | | |
| CMHS 7 | 0.77*** | 0.80 | 0.74 | 0.44 | 0.94 |
| RDS 3 | 0.69* | 0.85 | 0.53 | 0.32 | 0.93 |
| BJMHS 2 | 0.68* | 0.90 | 0.46 | 0.31 | 0.95 |
| Any current Axis I or Axis II psychiatric disorder (prevalence: 54.6%) | | | | | |
| CMHS 5 | 0.66** | 0.64 | 0.68 | 0.71 | 0.61 |
| RDS 2 | 0.61 | 0.76 | 0.46 | 0.63 | 0.61 |
| BJMHS 1 | 0.65* | 0.81 | 0.49 | 0.66 | 0.68 |
| Any current Axis I or Axis II disorder other than ASPD (prevalence: 34.0%) | | | | | |
| CMHS 6 | 0.80*** | 0.82 | 0.78 | 0.66 | 0.89 |
| RDS 2 | 0.67** | 0.88 | 0.45 | 0.45 | 0.88 |
| BJMHS 2 | 0.70** | 0.88 | 0.52 | 0.48 | 0.89 |
| **Black male inmate subsample** | | | | | |
| Affective disorder (current prevalence: 4.3%) | | | | | |
| CMHS 7 | 0.85* | 1.00 | 0.70 | 0.13 | 1.00 |
| RDS 9 | 0.81 | 0.67 | 0.96 | 0.40 | 0.98 |
| BJMHS 4 | 0.78 | 0.67 | 0.89 | 0.22 | 0.98 |
| Anxiety disorder (current prevalence: 9.4%) | | | | | |
| CMHS 4 | 0.71 | 1.00 | 0.41 | 0.15 | 1.00 |
| RDS 1 | 0.61 | 1.00 | 0.22 | 0.12 | 1.00 |
| BJMHS 1 | 0.66 | 1.00 | 0.32 | 0.13 | 1.00 |

**TABLE 4 (continued)**

| *Cut Point* | *AUC* | *Sensitivity* | *Specificity* | *PPP* | *NPP* |
|---|---|---|---|---|---|
| Any current Axis I disorder (prevalence: 10.1%) | | | | | |
| CMHS 6 | 0.74* | 0.86 | 0.61 | 0.20 | 0.97 |
| RDS 2 | 0.69 | 1.00 | 0.39 | 0.16 | 1.00 |
| BJMHS 5 | 0.63 | 0.29 | 0.97 | 0.50 | 0.92 |
| Any current Axis I or Axis II psychiatric disorder (prevalence: 50.7%) | | | | | |
| CMHS 7 | 0.68** | 0.51 | 0.85 | 0.78 | 0.63 |
| RDS 7 | 0.63 | 0.26 | 1.00 | 1.00 | 0.57 |
| BJMHS 1 | 0.66* | 0.86 | 0.47 | 0.63 | 0.76 |
| Any current Axis I or Axis II disorder other than ASPD (prevalence: 29.0%) | | | | | |
| CMHS 6 | 0.76** | 0.80 | 0.71 | 0.53 | 0.90 |
| RDS 2 | 0.64 | 0.85 | 0.43 | 0.38 | 0.88 |
| BJMHS 1 | 0.68* | 0.95 | 0.41 | 0.40 | 0.95 |

NOTE: $N = 201$ men. AUC = area under the curve; PPP = positive predictive power; NPP = negative predictive power; CMHS = Correctional Mental Health Screen; RDS = Referral Decision Scale; BJMHS = Brief Justice Mental Health Screen. Numbers following CMHS, RDS, and BJMHS refer to optimal cut point scores (Youden index). Predictive utility statistics for other cut points may be obtained from the first author by request.
*$p < .05$. **$p < .01$. ***$p < .001$.

significantly, correlated with a history of substance abuse treatment. All of the discriminant validity correlations were statistically significantly lower than the convergent validity correlations for the CMHS-F and the CMHS-M (Fischer's exact statistic for $z$-transformed $r \geq 1.88$, $p < .05$), except for the correlation between the CMHS-F and the suicidality risk rating. The CMHS-F's correlation with suicidality risk was significantly higher than that for the CMHS-F with the rating of vocational needs and the violence of criminal charges, but all other discriminant validity correlations for the CMHS-F were not significantly lower than the convergent validity correlation of the CMHS-F and suicidality.

With regard to the criterion validity of the screening measures, the mean CMHS-F scores differed for respondents who did and did not meet criteria for at least one current Axis I or Axis II psychiatric diagnosis ($M$ [$SD$] = 4.6 [2.2] vs. 2.45 [1.8], $t = 4.29$, $df = 98$, $p < .001$); a current Axis I or Axis II psychiatric diagnosis, excluding ASPD ($M$ [$SD$] = 4.8 [2.1] vs. 2.4 [1.9], $t = 5.2$, $df = 98$, $p < .001$); an affective disorder ($M$ [$SD$] = 5.1 [2.1] vs. 2.85 [2.0], $t = 5.4$, $df = 98$, $p < .001$); and an anxiety disorder ($M$ [$SD$] = 5.2 [2.0] vs. 2.95 [2.0], $t = 5.7$, $df = 95$, $p < .001$). CMHS-M scores differed for men who did and did not meet criteria for a current Axis I or II psychiatric diagnosis ($M$ [$SD$] = 5.8 [3.0] vs. 3.2 [2.4], $t = 6.15$, $df = 196$, $p < .001$); a psychiatric diagnosis, excluding ASPD ($M$ [$SD$] = 6.6 [2.9] vs. 3.4 [2.4], $t = 8.4$, $df = 196$, $p < .001$); an affective disorder ($M$ [$SD$] = 7.3 [2.5] vs. 4.1 [2.8], $t = 6.6$, $df = 195$ $p < .001$); and an anxiety disorder ($M$ [$SD$] = 6.5 [3.1] vs. 4.2 [2.9], $t = 4.7$, $df = 182$, $p < .001$).

## DISCUSSION

The CMHS-F and CMHS-M showed evidence of reliability, validity, and predictive utility for the identification of incarcerated women and men with undetected current Axis I or Axis II psychiatric disorders, except ASPD. Given the risk for new episodes and cumulative psychosocial impairment faced by adults with histories of psychiatric illness (Daley, Hammen, & Rao, 2000; Tyrer, Seivewright, & Johnson, 2004), particularly when confronted with the substantial adversities associated with criminal justice incarceration (Ditton, 1999), vulnerable persons who are not readily identified because they do not present with florid psychiatric symptoms or distress may be a critical subgroup of newly incarcerated adults to target for mental health screening. Future research is needed to determine if systematic evaluation and treatment following the early identification of these typically undetected at-risk men and women upon entry to jail can lead to improved correctional and postincarceration psychosocial and socioeconomic outcomes on both an individual and a systemic basis.

### Correctional Mental Health Screen–Female

The accuracy rate for the CMHS-F ranged from 70% to 80% for identifying undetected current Axis I or II psychiatric disorders among women, consistently exceeding the 61.6% accuracy reported with incarcerated women in the only other study of screening with incarcerated women (Steadman et al., 2005). In comparison with the

**FIGURE 3**
**Classification Accuracy for Identifying Female Inmates With Axis I or II Psychiatric Disorders (Excluding Antisocial Personality Disorder) of the Correctional Mental Health Scale–Female (CMHS-F), After Controlling for the Brief Jail Mental Health Screen (BJMHS; Solid Line), Compared With the BJMHS After Controlling for the CMHS-F (Dotted Line)**

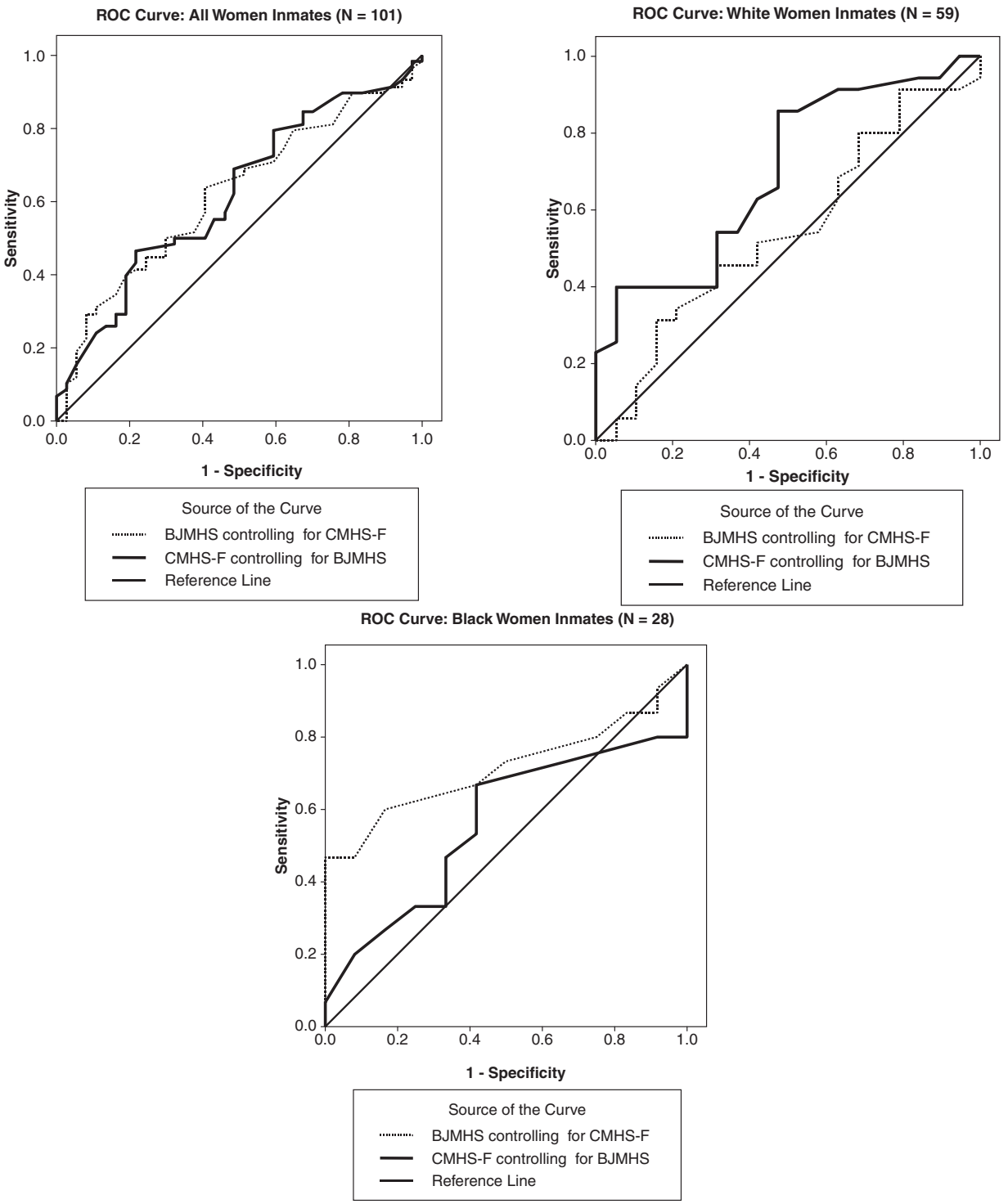
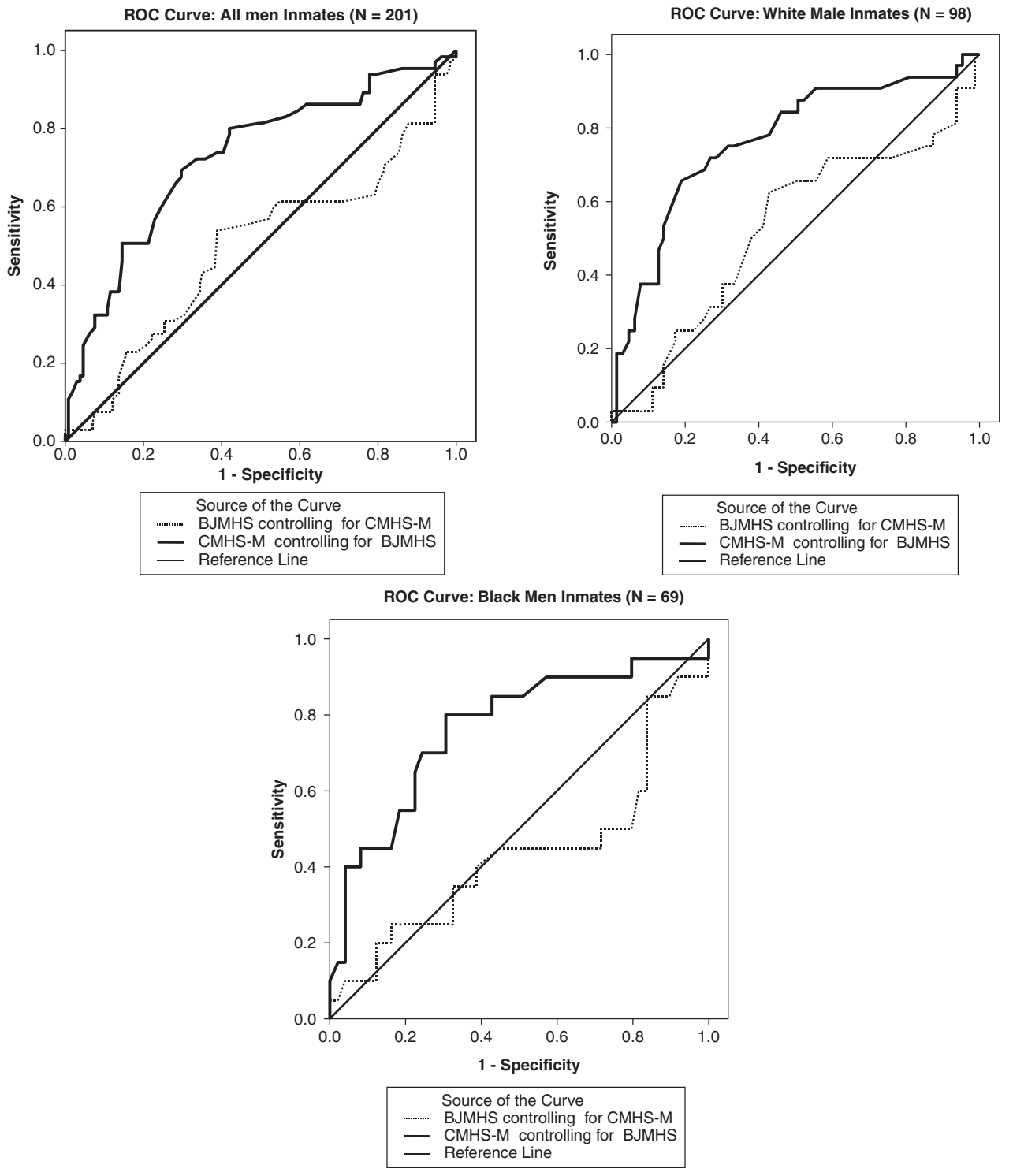


**FIGURE 4**
**Classification Accuracy for Identifying Male Inmates With Axis I or II Psychiatric Disorders (Excluding Antisocial Personality Disorder) of the Correctional Mental Health Scale–Male (CMHS-M) After Controlling for the Brief Jail Mental Health Screen (BJMHS; Solid Line) Compared With the BJMHS After Controlling for the CMHS-M (Dotted Line)**




NOTE: ROC = receiver-operating characteristic.

**TABLE 5**
**Pearson Correlations of CMHS-F and CMHS-M Scores With Correctional Variables**

| *Correctional Records Variable* | *CMHS-F* | *CMHS-M* |
|---|---|---|
| **Convergent validity** | | |
| Mental health score | .50** | .34** |
| Psychological evaluation conducted within 14 days | .43** | N/A |
| Suicide risk score | .14 | N/A |
| **Discriminant validity** | | |
| Overall behavioral risk score | .17 | .08 |
| Escape profile | .03 | .17 |
| Substance abuse treatment | .21* | .13 |
| Violence of current charges | –.08 | –.03 |
| Medical score | .19 | .05 |
| Disciplinary history | –.02 | .13 |
| Vocational training needs | –.09 | .00 |
| Educational training needs | .08 | –.12 |
| Sexual offender treatment | –.01 | .10 |

NOTE: CMHS-F = Correctional Mental Health Screen–Female; CMHS-M = Correctional Mental Health Screen–Male; N/A= records not available. $N = 101$ for CMHS-F and 201 for CMHS-M.
*$p < .05$ (two tailed). **$p < .01$.

35% false-negative rate reported by Steadman et al. (2005) with the BJMHS, the CMHS-F's median 21% false-negative rate appears to be an improvement. The 30% median false-positive rate of the CMHS-F also improved on the 45% false-positive rate reported for women with the BJMHS (Steadman et al., 2005).

In a direct test of the predictive utility of the CMHS-F with that of the BJMHS and the RDS (Teplin & Swartz, 1989), using cut points generated to optimize each screen's performance in this sample, the CMHS-F consistently had the highest overall accuracy, particularly in identifying White women with undetected current Axis I or II psychiatric disorders with or without ASPD (i.e., 9% to 12% greater AUCs). Thus, in contrast to studies that have not included jailed women (e.g., Teplin & Swartz, 1989), included female inmate samples too small to develop gender-specific screening instruments (Veysey et al., 1998), or used single screening measures for both genders that yielded problematic rates of false negatives for jailed women (Steadman et al., 2005), the CMHS-F shows promise as a mental health screen for incarcerated women in jails. In addition to its psychometric integrity and predictive utility, the CMHS-F is brief (eight items), identical in length to the shortest alternative screen, the BJMHS. When the low cut points identified in this study to maximize NPP are used, the CMHS-F shows promise as a Stage 1 screening tool for identifying clinically relevant psychiatric risk with a degree of precision that leaves few at-risk women undetected because of false negatives while not overwhelming the Stage 2 clinical mental health evaluation and treatment system with false positives. If minimizing false positives is a priority, the use of the higher optimal classification cut points with incarcerated White women may warrant consideration.

However, the CMHS-F was not more accurate than the BJMHS or the RDS overall with Black women, and it was less accurate than the BJMHS and RDS specifically in detecting Axis II disorders among Black women as a result of poorer specificity despite equivalent sensitivity. The CMHS-F showed poorer specificity than the RDS or the BJMHS also in identifying Axis I affective and anxiety disorders among Black women, although its sensitivity in identifying anxiety or affective disorders with Black women exceeded that of the RDS and BJMHS, leading to comparable overall accuracy. The higher rate of refusal to participate in the study by Black women may have led to an underrepresentation of Black women with psychiatric disorders, which could have undermined the ability of the CMHS-F to achieve specificity despite sensitively identifying Black women with undetected disorders. Future studies are needed to test the CMHS-F's utility with incarcerated Black women, both to examine and refine the screening criteria and to determine if altering the CMHS-F or the scoring rules for identifying cases can reduce the rate of false positives with Black women.

On the basis of these findings, further studies using the CMHS-F as a stand-alone brief screener (as opposed to being embedded in a longer composite screen, as in this initial study) are now needed to replicate and extend these initial promising results concerning its feasibility, reliability, validity, and predictive utility with newly incarcerated women. The correlation with receipt of a mental health evaluation within 14 days of incarceration suggests that the CMHS-F total score may serve as an early indicator of previously undetected mental health needs that emerge in the acute phase of incarceration. In addition, the statistically significant but small correlation between CMHS-F scores and a history of addiction treatment suggest that the CMHS-F may have utility for identifying women who are at risk for substance use problems as well as psychiatric risk or impairment. The finding that CMHS-F accuracy was poorer when ASPD was included suggests that the screen may primarily have utility for identifying psychiatric risk or psychosocial vulnerability that is distinct from ASPD. Finally, the finding that the CMHS-F showed strong convergent validity with institutional mental health needs ratings suggests that the CMHS-F may be particularly useful for identifying newly incarcerated women who are noted by correctional staff members as having mental health needs but who are not immediately viewed as having sufficient problems to warrant referral for services.

### Correctional Mental Health Screen–Male

The CMHS-M for men also showed evidence of reliability, convergent and discriminant validity, and predictive utility for the identification of undetected psychiatric disorders, except for ASPD. CMHS-M accuracy (75% to 80%) for identifying incarcerated men with undetected current psychiatric disorders, excluding ASPD, was superior, with both White and Black men, to that reported in a prior study with the BJMHS (73.5%; Steadman et al., 2005). Compared with the 15% false-negative rate reported for the BJMHS (Steadman et al., 2005) and the 5% to 13% rate found using optimal cut points with BJMHS items in the present study, the CMHS-M had a poorer (18% to 26%) false-negative rate for detecting an Axis I or II disorder, excluding ASPD. However, the 22% to 29% false-positive rate of the CMHS-M was about half the 49% false-positive rate reported for men with the BJMHS (Steadman et al., 2005) and the 48% to 59% false-positive rate found in this study using optimal cut points with the BJMHS. Thus, the CMHS-M may somewhat underdetect men with psychiatric morbidity, but it appears to better rule out male noncases than the BJMHS.

In a direct test of the predictive utility of the CMHS-M with that of the BJMHS and RDS using both the full range of scores and optimal classification cut points, the CMHS-M always exceeded or equaled the other screens' overall accuracy and was particularly strong in identifying the presence of affective disorders among White men and undetected Axis I or II psychiatric disorders, excluding ASPD, among Black men (i.e., 8% to 12% greater AUCs). In contrast to the comparison of screening instruments for women, the CMHS-M's advantage compared with the RDS and BJMHS held true for Black as well as White inmates and appeared to be primarily due to achieving superior specificity. With the exception of anxiety disorders, in the full sample and with White men particularly, the CMHS-M achieved lower rates of false positives (20% to 32%) than the RDS or the BJMHS (46% to 60%). Among Black men, the CMHS-M was more accurate overall, particularly for sensitivity in identifying affective disorders and specificity for detecting Axis I or Axis II disorders, excluding ASPD. All three screens performed most poorly in identifying anxiety disorders with men, perhaps in part because of an underreporting of anxiety symptoms by men (Simonds & Whiffen, 2003). CMHS-M scores were unrelated to history of past substance abuse treatment, suggesting that the CMHS-M specifically targets undetected mental health morbidity (as opposed to broader behavioral problems such as addiction).

As was true with the CMHS-F, the CMHS-M did not perform as well when the criterion psychiatric disorders included ASPD as when ASPD was excluded. The CMHS was not designed to identify ASPD: No items reflecting ASPD were included in the instruments assembled for the composite screening measure. Thus, the poorer predictive accuracy documented when ASPD was included as a criterion diagnosis in effect supports the discriminant validity of the CMHS. The CMHS-M (and CMHS-F) showed evidence of predictive utility when Axis II disorders other than ASPD were included in the criterion set of diagnoses as well as Axis I psychiatric disorders. This finding suggests that the CMHS may have utility in identifying Axis II disorders such as Cluster A paranoid, schizoid, and schizotypal disorders or Cluster B borderline personality disorder. These Axis II disorders often are comorbid with ASPD and Axis I disorders and may account for much of the functional impairment and behavior management problems that are typically attributed to more readily recognized disorders such as ASPD (Coid, 2003). Inmates who are floridly psychotic, agitated, or antisocial are routinely identified and triaged for specialized evaluation and supportive custody placement. However, inmates with undetected personality pathology (Tyrer et al., 2004) may develop severe functional impairment and behavior management problems. Further testing of the utility of the CMHS-M for identifying potentially costly and dangerous Axis II disorders other than the more readily identified problems of ASPD therefore appears warranted with newly incarcerated men.

### Limitations

This study's findings should be interpreted in light of limitations in methodology. The composite screen items may not reflect the full domain of psychiatric illness (Zimmerman & Mattia, 2001). The brief screens also were not freestanding but were embedded in the composite screen. Cross-validation of feasibility and predictive utility is needed with the CMHS-F and CMHS-M administered on a stand-alone basis. Tests of incremental utility also require head-to-head comparisons of the BJMHS and RDS when those screens are administered on a stand-alone basis (rather than with the items embedded in a composite screen, as was done in this study). In such studies, item analyses will be important to examine the extent to which each item contributes to predictive utility, as well as if different scoring parameters (e.g., item weights, scoring rules) can enhance utility.

The presence of psychiatric morbidity in this sample may be understated as a result of factors that militate against recognizing or acknowledging clinically significant psychiatric symptoms during the stressful transition from arrest to initial jail incarceration (Gavin et al., 2003), so screening based on current symptoms or disorders at jail admission may have limited utility in predicting subsequent

need for mental health services. However, most adults with psychiatric disorders at some point in their lifetimes do seek treatment, but typically only after delays that often are as long as several decades (Wang et al., 2005). Therefore, identifying inmates who are willing to report clinically significant problems upon entry to jail may provide a basis for enhanced surveillance and preventive services during subsequent high-risk periods (e.g., if sentenced to prolonged incarceration, at community reentry). Studies will be needed to determine if the CMHS can identify inmates who subsequently develop high acuity or severe impairment because of mental health problems and to test different treatments to develop cost-effective models for the management of subclinical impairment and the secondary prevention of severe impairment.

The study's findings also are limited in generalizability to jails (compared with other correctional facilities, such as prisons and forensic hospitals), to newly incarcerated inmates, and to inmates not previously identified with mental health needs. We focused on undetected cases because these are the persons for whom additional screening will not be redundant with the procedures already in place in many correctional programs, but as a result, our findings may underestimate the utility of the CMHS. The underrepresentation of Black and Latino individuals because of refusals and language barriers detracts from the otherwise representative sampling of the target population and suggests a need for further tests of the CMHS with minorities. The study also excluded inmates with known mental illnesses to focus on the identification of inmates with otherwise undetected psychiatric risk or impairment.

With these limitations in mind, the study results suggest that the CMHS-F and CMHS-M warrant further testing as potentially practical, reliable, valid, and accurate tools for the identification of newly incarcerated men and women with, or at risk for, undetected psychiatric impairment. The gender-specific and ethnicity-specific item content and cut points identified for the CMHS-F and CMHS-M will facilitate the accurate screening of both genders and White and Black inmates. On the basis of both a comparison with prior studies' findings and a direct comparison using optimal cut points for the present sample, the CMHS was more accurate than two well-validated alternative screening instruments in identifying jail detainees who were undetected by the correctional system but who met criteria for current psychiatric morbidity on independent research diagnostic testing. The CMHS thus has promise as an empirically based gender- and ethnicity-specific screening measure for use in detecting previously undetected psychiatric morbidity among newly incarcerated adult jail detainees.

## REFERENCES


Abram, K., Teplin, L., & McLelland, G. (2003). Comorbidity of severe psychiatric disorders and substance use disorders among women in jail. *American Journal of Psychiatry*, *160*, 1007-1010.

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

Andersen, H. S., Sestoft, D., Lillebaek, T., Gabrielson, G., & Hemmingsen, R. (2002). Validity of the General Health Questionnaire (GHQ-28) in a prison population: Data from a randomized sample of prisoners on remand. *International Journal of Law and Psychiatry*, *25*, 573-580.

Beck, A. J., & Maruschak, L. M. (2001, July). Mental health treatment in state prisons. *Bureau of Justice Statistics Special Report*, pp. 1-8.

Book, A., Knap, M., & Holden, R. (2001). Criterion validity of the Holden Psychological Screening Inventory Social Symptomatology Scale in a prison sample. *Psychological Assessment*, *13*, 249-253.

Boothby, J. L., & Durham, T. W. (1999). Screening for depression in prisoners using the Beck Depression Inventory. *Criminal Justice and Behavior*, *26*, 107-124.

Boutin-Foster, C., Ferrando, S., & Charlson, M. (2003). The Cornell Psychiatric Screen: A brief psychiatric scale for hospitalized medical patients. *Psychosomatics*, *44*, 382-387.

Brinded, P., Simpson, A., Laidlaw, T., Fairley, N., & Malcolm, F. (2001). Prevalence of psychiatric disorders in New Zealand prisons. *Australian and New Zealand Journal of Psychiatry*, *35*, 166-173.

Cerny, B. A., & Kaiser, F. H. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioral Research*, *12*, 43-47.

Coid, J. (2003). The co-morbidity of personality disorder and lifetime clinical syndromes in dangerous offenders. *Journal of Forensic Psychiatry and Psychology*, *14*, 341-366.

Connecticut Department of Correction. (2005). *Connecticut Department of Correction objective classification manual* (rev. 7/05). Available at http://www.ct.gov/doc/lib/doc/PDF/PDFReport/ClassificationManualLibraryCopy.pdf

Cooke, D. J. (1998). The development of the Prison Behaviour Rating Scale. *Criminal Justice and Behavior*, *25*, 482-506.

Daley, S., Hammen, C., & Rao, U. (2000). Predictors of first onset and recurrence of major depression in young women during the 5 years following high school graduation. *Journal of Abnormal Psychology*, *109*, 525-533.

Daradkeh, T. K., Ghubash, R., El-Rufaie, O.E.F., & Abou-Saleh, M. T. (1999). The rationale, development and reliability of a new screening psychiatric instrument. *Social Psychiatry and Psychiatric Epidemiology*, *34*, 223-228.

Diamond, P., Wang, E., Holzer, C., III, Thomas, C., & Cruser, A. (2001). The prevalence of mental illness in prison. *Administration and Policy in Mental Health*, *29*, 21-40.

DiCataldo, F., Greer, A., & Profit, W. E. (1995). Screening prison inmates for mental disorder. *Bulletin of the American Academy of Psychiatry and the Law*, *23*, 573-585.

Ditton, P. M. (1999, July). Mental health treatment of inmates and probationers. *Bureau of Justice Statistics Special Report*, pp. 1-12.

Edens, J. F., Skeem, J. L., & Douglas, K. S. (2006). Incremental validity analyses of the Violence Risk Appraisal Guide and the Psychopathy Checklist: Screening version in a civil psychiatric sample. *Assessment*, *13*, 368-374.

Ericsson, M., Poston, W., Linder, J., Taylor, J., Haddock, C. K., & Foreyt, J. (2002). Depression predicts disability in long-term chronic pain patients. *Disability and Rehabilitation*, *24*, 334-340.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J.B.W. (1996). *Structured Clinical Interview for Axis I and II DSM-IV Disorders–Patient Edition (SCID-IV/P)*. New York: New York State Psychiatric Institute, Biometrics Research Department.

Furukawa, T. A., Goldberg, D., Rabe-Hesketh, S., & Uestuen, T. (2001). Stratum-specific likelihood ratios of two versions of the General Health Questionnaire. *Psychological Medicine*, *31*, 519-529.

Gagliardi, G., Lovell, D., Peterson, P., & Jemelka, R. (2004). Forecasting recidivism in mentally ill offenders released from prison. *Law & Human Behavior*, *28*, 133-155.

Gavin, N., Parsons, S., & Grubin, D. (2003). Reception screening and mental health needs assessment in a male remand prison. *Psychiatric Bulletin*, *27*, 251-253.

Goodman, L., Corcoran, C., Turner, K., Yuan, N., & Green, B. L. (1998). Assessing traumatic event exposure: General issues and preliminary findings for the Stressful Life Events Screening Questionnaire. *Journal of Traumatic Stress*, *11*, 521-542.

Hart, S., Roesch, R., Corrado, R., & Cox, D. (1993). The Referral Decision Scale: A validation study. *Law and Human Behavior*, *1*, 611-623.

Herrmann, C. (1997). International experiences with the Hospital Anxiety and Depression Scale. *Journal of Psychosomatic Research*, *42*, 17-41.

Jordan, B. K., Federman, E. B., Burns, B., Schlenger, W., Fairbank, J., & Caddell, J. (2002). Lifetime use of mental health and substance abuse treatment services by incarcerated women felons. *Psychiatric Services*, *53*, 317-325.

Jordan, B. K., Schlenger, W., Fairbank, J., & Cadell, J. (1996). Prevalence of psychiatric disorders among incarcerated women, II: Convicted felons entering prison. *Archives of General Psychiatry*, *53*, 513-519.

Kessel, J. B., & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires. *Psychological Assessment*, *5*, 395-399.

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Moroczek, D., Normand, S., et al. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, *32*, 959-976.

Langbehn, D., Pfohl, B., Reynolds, S., Clark, L., Battagila, M., Bellodi, L., et al. (1999). The Iowa Personality Disorders Screen. *Journal of Personality Disorders*, *13*, 75-89.

McLearen, A., & Ryba, N. (2003). Identifying severely mentally ill inmates. *Journal of Offender Rehabilitation*, *37*, 25-40.

Metzner, J., Miller, R. D., & Kleinsasser, D. (1994). Mental health screening and evaluation within prisons. *Bulletin of the American Academy of Psychiatry and the Law*, *22*, 451-457.

Parsons, S., Walker, L., & Grubin, D. (2001). Prevalence of mental disorder in female remand prisons. *Journal of Forensic Psychiatry*, *12*, 194-202.

Perkins, N., & Schisterman, E. (2006). The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *American Journal of Epidemiology*, *163*, 670-675.

Prins, A., Ouimette, P., Kimerling, R., Cameron, R., Hugelshofer, D., Shaw-Hegwer, J., et al. (2003). The Primary Care PTSD Screen (PC-PTSD): Development and operating characteristics. *Primary Care Psychiatry*, *9*, 9-14.

Resnick, H. S., Falsetti, S. A., Kilpatrick, D. G., & Freedy, J. R. (1996). Assessment of rape and other civilian trauma-related post-traumatic stress disorder. In T. W. Miller (Ed.), *Stressful life events* (pp. 231-266). Madison, WI: International Universities Press.

Rogers, R., Sewell, K., Ustad, K., Reinhardt, V., & Edwards, W. (1995). The Referral Decision Scale with mentally disordered inmates. *Law and Human Behavior*, *19*, 481-492.

Shrout, P., Skodol, A., & Dohrenwend, B. P. (1986). A two stage approach for case identification and diagnosis: First stage instruments. In J. Barrett & R. Rose (Eds.), *Mental disorder in the community* (pp. 286-303). New York: Guilford.

Simonds, V., & Whiffen, V. (2003). Are gender differences in depression explained by gender differences in co-morbid anxiety? *Journal of Affective Disorders*, *77*, 197-202.

Smith, C., & Borland, J. (1999). Minor psychiatric disturbance in women serving a prison sentence. *Legal and Criminological Psychology*, *4*, 273-284.

Spitzer, R., Kroenke, K., & Williams, J. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ Primary Care Study. *JAMA*, *282*, 1737-1744.

Steadman, H., Scott, J., Osher, F., Agnese, T., & Robbins, P. (2005). Validation of the Brief Jail Mental Health Screen. *Psychiatric Services*, *56*, 816-822.

Teplin, L. (1990). The prevalence of severe mental disorder among male urban jail detainees. *American Journal of Public Health*, *80*, 663-669.

Teplin, L. (1994). Psychiatric and substance abuse disorders among male urban jail detainees. *American Journal of Public Health*, *84*, 290-293.

Teplin, L., Abram, K., & McLelland, G. (1996). Prevalence of psychiatric disorders among incarcerated women: Pretrial jail detainees. *Archives of General Psychiatry*, *53*, 505-512.

Teplin, L., Abram, K., & McClelland, G. (1997). Mentally disordered women in jail: Who receives services? *American Journal of Public Health*, *87*, 604-609.

Teplin, L., & Swartz, J. (1989). Screening for severe mental disorder in jails: The development of the Referral Decision Scale. *Law and Human Behavior*, *13*, 1-18.

Tyrer, P., Seivewright, H., & Johnson, T. (2004). The Nottingham Study of Neurotic Disorder. *Psychological Medicine*, *34*, 1385-1394.

Veysey, B., Steadman, H., Morrissey, J., Johansen, M., & Beckstead, J. (1998). Using the Referral Decision Scale to screen mentally ill jail detainees. *Law and Human Behavior*, *22*, 205-215.

Walters, G. D., & Chlumsky, M. L. (1993). The Lifestyle Criminality Screening Form and ASPD. *Behavioral Sciences and the Law*, *11*, 111-115.

Wang, P., Berglund, P., Olfson, M., Pincus, H., Wells, K., & Kessler, R. (2005). Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*, 603-613.

Weathers, F., Ruscio, A., & Keane, T. (1999). Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychological Assessment*, *11*, 124-133.

Zimmerman, M., & Mattia, J. I. (2001). The Psychiatric Diagnostic Screening Questionnaire: Development, reliability and validity. *Comprehensive Psychiatry*, *42*, 175-189.