

# **COMPAS Validation Study: Final Report**

*August 15, 2010*

**Prepared by  
David Farabee (UCLA)  
Sheldon Zhang (SDSU)  
Robert E.L. Roberts (CSUSM)  
Joy Yang (UCLA)**

Semel Institute for Neuroscience and Human Behavior  
University of California, Los Angeles  
1640 S. Sepulveda, Suite 200  
Los Angeles, California 90025

*California Department of Corrections and Rehabilitation  
Agreement No: C06.417*



## EXECUTIVE SUMMARY

COMPAS (Correctional Offender Management and Profiling Alternative Sanctions) is a computerized database and analysis system designed to help criminal justice practitioners determine the placement, supervision, and case-management of offenders in community and secure settings. The California Department of Corrections and Rehabilitation contracted with the University of California, Los Angeles (UCLA), and San Diego State University (SDSU) to validate the instrument in terms of its ability to identify treatment needs among inmates as well as predict various recidivism outcomes.

A total of 91,334 parolees who had been assessed with COMPAS prior to release were included in the study sample. Of these, roughly 60,000 had been on parole for at least 12 months and the remainder had been on parole for at least 24 months. Characteristics of the study subjects closely paralleled those of the general parolee population in California.

### Validation of Needs Scales

The COMPAS needs scales were evaluated in terms of their reliability over time (test-retest coefficients) and the extent to which their constituent scales correlated with relevant counterparts on the Level of Service Inventory-Revised (LSI-R) scale (concurrent validity). To accomplish this, the COMPAS was administered twice to 75 inmates at the California Institute for Men (CIM) located in Chino, California, over a span of approximately two weeks. To establish concurrent validity, the LSI-R was also administered at the same time points.

The COMPAS scales showed extremely high test-retest reliability, ranging from .70 to 1.00. The perfect and near-perfect correlations obtained for many of the scales appear to be driven by the fact that these scales were coded directly from the inmates' Central Files. Overall, the average test-retest correlation coefficient for the COMPAS scales was .88.

Of the 18 scales making up the core of the COMPAS assessment, nine appeared to measure identical or similar constructs with scales found in the LSI-R. For six of these scales (Criminal Involvement, Criminal Associates/Peers, Substance Abuse, Financial, Vocational/Educational, and Housing), significant and positive correlations were found between the COMPAS and LSI-R. The correlations were marginally significant for two of the scales, Family Criminality (COMPAS) with Family/Marital (LSI-R) and Criminal Attitudes (COMPAS) with Attitudes/Orientation (LSI-R), and not significant for one, Leisure (COMPAS) with Leisure/Recreation (LSI-R).

### Validation of Risk Scales

Using official records data provided by the California Department of Corrections and Rehabilitation (CDCR), two major outcome measures were examined: (1) any subsequent arrest, and (2) a subsequent arrest for a violent offense. For the first measure, the overall

# COMPAS Final Report | 2010

---

## TABLE OF CONTENTS

PREFACE .....	2
EXECUTIVE SUMMARY .....	3
I. BACKGROUND .....	8
II. METHODS .....	8
II.A. Validation of COMPAS Needs Assessment Components.....	8
II.B. Validation of COMPAS Risk Scales .....	9
II.B.1. Data Sources .....	10
II.B.2. Sample.....	10
III. IMPLEMENTATION OF COMPAS.....	10
III.A. Parolee Services Administrator Survey.....	11
III.B. Perceived Benefits from PSA Interviews and Field Observations.....	12
III.C. Perceived Problems from PSA Interviews and Field Observations.....	12
III.D. Expanded Use of COMPAS in California .....	13
IV. ANALYSIS AND FINDINGS.....	13
IV.A. Test-Retest Reliability of Needs Scales.....	13
IV.B. Concurrent Validity of COMPAS Needs Scales with the LSI-R .....	15
IV.B.1. Criminal Involvement.....	16
IV.B.2. Criminal Peers .....	16
IV.B.3. Substance Abuse.....	17
IV.B.4. Financial .....	17
IV.B.5. Vocational/Education .....	18
IV.B.6. Family.....	18
IV.B.7. Leisure/Recreation .....	19
IV.B.8. Housing.....	19
IV.B.9. Criminal Thinking .....	20
IV.C. Predictive Validity of Risk Scales.....	20
IV.C.1. Descriptive Statistics.....	21
IV.C.2. Receiver Operating Characteristics .....	24
IV.C.3. Logistic Regression and Comparative Predictive Power Analysis .....	27
V. SUMMARY AND DISCUSSION.....	28
APPENDIX .....	30

## LIST OF FIGURES

Figure 1: Percentage of Inmates Categorized as Having High Levels of Criminal Involvement .....	16
Figure 2: Percentage of Inmates Categorized as having High Problem Levels of Criminal Peers/Associates .....	16
Figure 3: Percentage of Inmates Categorized as "High" on Substance Abuse Problems ...	17
Figure 4: Percentage of Inmates Categorized as "High" on Financial Problems.....	17
Figure 5: Percentage of Inmates Categorized as Being in High Need of Vocational/Educational Assistance .....	18
Figure 6: Percentage of Inmates Categorized as Having High Levels of Family Problems .	19
Figure 7: Percentage of Inmates Categorized as Having High Need for Assistance in Leisure/Recreation .....	19
Figure 8: Percentage of Inmates Categorized as Having High Need for Housing Assistance .....	20
Figure 9: Percentage of Inmates Categorized as Having High Levels of Criminal Thinking	20
Figure 10: Percentage of Parolees Arrested within Two Years of Parole Release by COMPAS Recidivism Score Decile .....	25
Figure 11: ROC Chart of COMPAS Recidivism Risk Scale Score Decile (Outcome: Arrested within Two Years of Parole Release) .....	25
Figure 12: Percentage of Parolees Arrested for Violent Offense within Two Years of Release by COMPAS Violent Recidivism Risk Score Decile .....	26
Figure 13: ROC Chart of COMPAS Violent Recidivism Risk Score Decile (Outcome: Arrested for Violent Offense within Two Years of Parole Release) .....	26
Figure 14: Percentage of Parolees Returned To Custody within Two Years by COMPAS Recidivism Risk Score Decile .....	30
Figure 15: ROC Chart of COMPAS Violent Recidivism Risk Scale Score Decile (Outcome: Violent Parole Violation within Two Years of Parole Release). .....	30
Figure 16: Percentage of Parolees that Committed A Violent Parole Violation within Two Years of Parole Release by COMPAS Violent Recidivism Risk Score Decile.....	31
Figure 17: ROC Chart of COMPAS Violent Recidivism Risk Scale Score Decile (Outcome: Violent Parole Violation within Two Years of Parole Release) .....	31

assessment were compared to scores obtained on its counterpart scales on the COMPAS. The LSI-R was also administered twice to demonstrate the consistency and reliability of the instrument for the same population.

The procedures for this component of the evaluation were as follows:

1. UCLA researchers contacted CIM-based Parole Service Associates (PSAs) ahead of time to schedule interview visits.
2. When UCLA researchers arrived at CIM, they met PSAs in their office and reviewed their Central Files (or "C-files") there.
3. Inmates who had not yet been given a COMPAS assessment and who had a release date at least 12 weeks away were randomly selected to participate in the study. The names and CDCR numbers of the selected participants were made available to UCLA researchers, who assigned them a unique study ID number.
4. On the West (Level 1) Yard, potential subjects were ducated in groups of about 10 and invited to participate, using the study recruitment script. If inmates agreed to participate, they were consented at this time. On subsequent visits, these subjects were ducated two at a time. Before they were ducated, a classroom, or other space was reserved for the interviews. One PSA and one UCLA researcher shared a classroom whenever possible to interview participants. They conducted the interviews on opposite sides of the room and did so in a way that preserved the confidentiality of the participants. The PSAs administered the COMPAS and the UCLA researchers administered the LSI-R. After the inmates were administered the first assessment, they switched seats and proceeded to take the remaining assessment. This counterbalanced design mitigated any potential ordering effects.
5. After the first administration of the COMPAS and LSI-R (Test), the UCLA researchers scheduled the clients for their follow-up (Retest), approximately 10 days later.
6. Money orders in the amount of \$10 were deposited into the subjects' accounts within one week of the test/retest.

These data were collected between March 2 and June 22, 2009.

### **II.B. Validation of COMPAS Risk Scales**

The research team evaluated the degree to which the COMPAS risk scores predicted future recidivism among California parolees. Two key outcomes were used to assess the predictive power of COMPAS—(1) a subsequent arrest for any reason following release, and (2) a subsequent arrest for a violent offense, including homicide, assault, sexual assault, robbery, domestic violence, and kidnapping.

sections provide survey and interview data that describe PSAs' perceptions of the COMPAS.

### III.A. Parolee Services Administrator Survey

To assess PSAs' perceptions of the COMPAS, UCLA staff e-mailed the survey to the Parole Planning and Placement Supervisor, who printed copies for the seven PSAs who administered the COMPAS at the California Men's Institution (CIM).<sup>3</sup> A large manila envelope was placed in a location in the Parole Planning and Placement office where PSAs could confidentially insert their completed surveys. All seven PSAs responded to the survey.

All of the respondents reported that the COMPAS is easy to use (mean = 1.3 [SD = 0.5] on a scale of 1 = "very easy" to 10 = "very difficult") and all reported having received COMPAS training before administering the COMPAS to inmates. The average time PSAs reported spending in training was 10.6 hours (SD = 5.5). The PSAs generally believed that the training prepared them well to administer the COMPAS (mean = 8.9 [SD = 1.2] on a scale from 1 = "not at all" to 10 = "very well"). One respondent noted that in addition to the initial COMPAS training, ongoing COMPAS training was also provided as issues and needs arose.

Each PSA administered the COMPAS Re-entry an average of 67 times (SD = 42.7) each month, and spent approximately 39 minutes (SD = 12.1) on each administration of the COMPAS Re-entry assessment interview. In addition, PSAs reported spending an average of 58 minutes (SD = 23.2) reviewing each inmate's C-file and, on average, 24 minutes (SD = 16.0) entering COMPAS data into the database system.

On average, the PSAs reported experiencing technical difficulties with the COMPAS data system 1.4 times (SD = 0.5) each month. They reported not being able to access the COMPAS system to either retrieve or enter data 33 minutes (SD = 26.9) per month, on average.

Respondents indicated that virtually all (97.6%) of inmates on their caseloads were given a list of referrals to various services (including employment, educational, transitional housing, and substance abuse) before they were released from prison.

When asked how utilization of the COMPAS might be improved, two respondents wrote that the COMPAS would be more effective if it could be used by custody staff (including R&R, Control, and Visiting Sergeants/Lieutenants) in the prisons. One felt this should be done for safety/security purposes and the other felt that the data would help the custody staff understand each inmate better.

---

<sup>3</sup> Due to its proximity to Los Angeles, CIM served as a "focus" site for all field data collection.

### III.D. Expanded Use of COMPAS in California

COMPAS has become the instrument of choice of CDCR major divisions as well as by an increasing number of community-based correctional service providers in California. The increasing popularity of the COMPAS over other risk and needs instruments (e.g., Level of Service Inventory-Revised) by community service providers in California is a recent development. Much of the impetus behind the adoption of COMPAS by community agencies is due to the requirement by Parole to base correctional services on systematic client assessment. According to service providers, cost is a major factor in their decision to adopt COMPAS. The COMPAS costs less than the more widely used LSI-R.

But the proximal impact of the COMPAS could not be determined in the current study. Specifically, the extent to which the assessment information generated by the instrument has been used by parole agents in the field remains unclear. In our limited interactions with field agents who have seen COMPAS reports, responses were less sanguine than those of the PSAs regarding the utility of the instrument. Of those who expressed somewhat positive attitudes toward the instrument, a common response was that the information generated by COMPAS validated their own assessments and provided greater confidence in their supervision and programming plans.

## IV. ANALYSIS AND FINDINGS

This section summarizes the results of the two quantitative components of this evaluation: (1) the validation of the COMPAS needs scales (including both test-retest and concurrent validity with the LSI-R), and (2) an assessment of the predictive validity of the COMPAS risk scales.

### IV.A. Test-Retest Reliability of Needs Scales

Test-retest coefficients range from 0 to 1, with higher values indicating greater stability of responses over time. Although high test-retest reliability alone does not constitute proof that a scale is valid, it is important to note that a test cannot be considered valid unless it demonstrates acceptable test-retest reliability.

As seen in Table 1, the COMPAS scales showed high test-retest reliability, ranging from .70 to 1.00. The perfect and near-perfect correlations obtained for many of the scales appear to be driven by the fact that these scales were coded directly from the inmates' Central Files. However, even the scales derived from inmates' self-reports demonstrated high reliability over time. Overall, the average test-retest correlation coefficient for the COMPAS scales was .88; for LSI-R, it was .64.

#### IV.B. Concurrent Validity of COMPAS Needs Scales with the LSI-R

Of the 18 scales making up the core of the COMPAS assessment, nine appeared to measure identical or similar constructs with the Level of Service Inventory-Revised (LSI-R), a widely recognized commercial risk/needs assessment instrument developed by Andrews and Bonta (2001). These counterpart scales are shown in Table 3 below.

Table 3: *Overlapping Constructs of the COMPAS and LSI-R Assessments*

COMPAS	LSI-R
Criminal Involvement	Criminal History
History of Non Compliance	
History of Violence	
Current Violence	
Criminal Associates/Peers	Companions
Substance Abuse	Alcohol/Drug Problems
Financial Problems/Poverty	Financial
Vocational/Education Problems	Education/Employment
Criminal Thinking	Attitudes/Orientation
Family Criminality	Family/Marital
Social Environment Problems	
Leisure and Recreation	Leisure/Recreation
Residential Instability	Accommodation
Social Adjustment Problems	
Socialization Failure	
Criminal Opportunity	
Criminal Personality	
Social Isolation	Emotional/Personal

Comparisons of the COMPAS and LSI-R scales were made by (1) correlating the continuous scales scores, and (2) by examining the concordance in determining whether a problem exists. These categories were based on cutoff scores provided by Northpointe Institute for Public Management and Multi-Health Systems, Inc. Briefly, for all of the COMPAS scales except Substance Abuse, high problem levels were defined as having a decile score equal to or greater than 8. For the Substance Abuse scale, the threshold was a decile score of at least 5. For the LSI-R scales, scores on the constituent items of the scale must first be summed, divided by the number of items, then multiplied by 4. This resulted in a 5-level problem indicator ranging from 0 (very low) to 4 (very high). For purposes of the current study (and consistent with the LSI-R scoring guide), scores of 2.5 or higher were categorized as high need (i.e., the "high" and "very high" groups were combined).

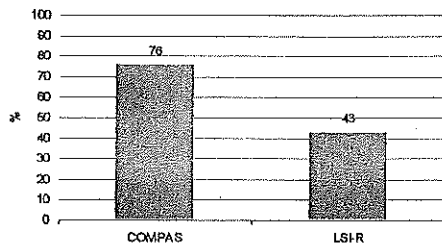


scoring high on the Companions scale, 50% also scored high on the CASSPEER. The overall phi co-efficient was .08.

#### IV.B.3. Substance Abuse

Over three quarters of the sample were designated as having a substance abuse problem, based on the COMPAS measure; 43% were so designated according to the LSI-R.

Figure 3: Percentage of Inmates Categorized as "High" on Substance Abuse Problems



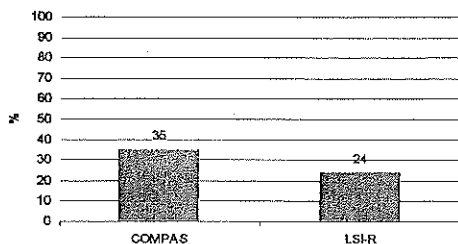
The correlation between the Substance Abuse scale (SUBABUSE) on the COMPAS and the Alcohol/Drug Problem scale on the LSI-R was .53 ( $p < .0001$ ). With regard to concordance, 54% of those scoring high on the SUBABUSE scale also scored high on the Alcohol/Drug Problem scale. Fully 97% of those identified as having a substance use problem on the LSI-R were also identified as having a problem on the COMPAS, suggesting that the LSI-R tended to render a more conservative estimate on drug/alcohol problems. The overall phi co-efficient between these two measures was .42.

An additional *yes/no* question was included in the interview for which inmates were asked whether they felt that they were in need of substance abuse treatment. It is important to note that this additional question was intended to reflect treatment need, rather than the mere acknowledgment of a problem. The phi coefficients between this treatment need item and the COMPAS and LSI-R substance abuse problem scales were .60 and .37, respectively.

#### IV.B.4. Financial

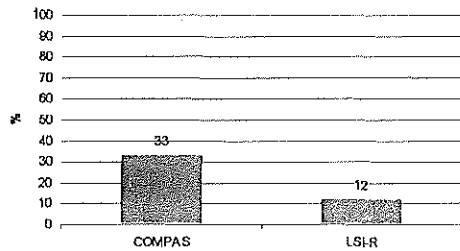
According to the COMPAS, slightly over one-third of the sample were experiencing significant financial problems. The LSI-R generated a more conservative estimate of 24%.

Figure 4: Percentage of Inmates Categorized as "High" on Financial Problems



of those scoring high on the FAMCRIM scale also scored high on the Family/Marital scale. Conversely, of those scoring high on the Family/Marital scale, 44% also scored high on the FAMCRIM scale. The overall phi co-efficient was .08.

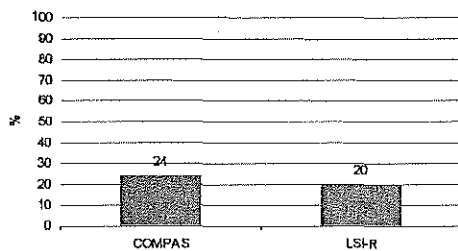
Figure 6: Percentage of Inmates Categorized as Having High Levels of Family Problems



#### IV.B.7. Leisure/Recreation

The LSI-R (20%) and COMPAS (24%) showed a similar prevalence of problems on the Leisure/Recreation variable. However, the agreement between these measures was quite low.

Figure 7: Percentage of Inmates Categorized as Having High Need for Assistance in Leisure/Recreation



The correlation between the LEISURE scale (COMPAS) and the Leisure/Recreation scale (LSI-R) was .05 ( $p > .10$ ). With regard to concordance, 28% of those scoring high on the COMPAS scale also scored high on the LSI-R counterpart scale. Of those scoring high on the Criminal History scale, 33% also scored high on the Criminal Involvement scale. The overall phi co-efficient was .11.

#### IV.B.8. Housing

The Residential Instability (RESINST) scale on the COMPAS and the Accommodation scale on the LSI-R suggested high levels of need for housing assistance, 39% and 25% of the sample, respectively.

The correlation between the RESINST and the LSI-R's Accommodation scale was .57 ( $p < .0001$ ). With regard to concordance, 52% of those scoring high on the COMPAS RESINST scale also scored high on the Accommodation scale. Of those scoring high on the Accommodation scale, 79% also scored high on the RESINST scale. The overall phi co-efficient was .48.

# COMPAS Final Report | 2010

## IV.C.1. Descriptive Statistics

Table 4 presents the background characteristics of the entire COMPAS sample, and subsamples with one- and two-year post-release observation times. For comparison purposes, the table also includes the population of prisoners released to parole during 2008. The COMPAS samples had a slight overrepresentation of African American and White parolees, as well as parolees whose most recent incarceration had been for a crime against a person. The COMPAS parolees were also slightly younger in the aggregate than the overall parole population.

Table 4: *Demographic Characteristics of Parolee Population, COMPAS Sample and Subsamples with Increasing Post-Release Observation Periods*

	Parole Population (2008) Percent	COMPAS Sample* Percent	COMPAS Sample with at Least One Year Observation Period Percent	COMPAS Sample with at Least Two Years Observation Period Percent
Gender				
Female	10.9	10.1	11.0	11.4
Male	89.1	89.9	89.0	88.6
Race/Ethnicity				
African American	24.0	27.1	26.5	26.9
Latino	41.1	37.0	37.5	37.0
White	29.8	30.9	31.1	31.5
Other	5.1	5.0	4.9	4.6
Age at Release				
Median	37.0	33.7	33.4	33.2
Principal Commitment Offense				
Property	29.4	25.8	24.5	23.7
Persons	27.0	31.8	32.6	33.6
Drugs	29.9	29.6	30.0	30.6
Other	13.7	12.8	12.9	12.2
Year of COMPAS Assessment				
2006	n/a	22.6	33.8	78.1
2007	n/a	40.8	58.6	21.9
2008	n/a	35.0	7.5	0.0
2009	n/a	1.6	0.0	0.0
Number of Parolees	123,665	91,334	60,793	25,009

Note: Sample includes those parolees with arrest history data and COMPAS scores for violent recidivism risk.

## COMPAS Final Report | 2010

Table 6: *Spearman Rank-Order Correlations between Two COMPAS Recidivism Risk Scale Scores (Deciles) and Arrest Type within Two Years of Release (25,009 Parolees with Two Year Observation Period).*

	Recidivism Risk Score Decile	Violent Recidivism Risk Score	Any Arrest within Two Years	Arrest for Violent Offense within Two Years	Arrest for Violent Felony within Two Years
Recidivism Risk Score Decile	1.00				
Violent Recidivism Risk Score Decile	0.68	1.00			
Any Arrest in Two Years	0.31	0.29	1.00		
Violent Arrest in Two Years	0.19	0.21	0.34	1.00	
Violent Felony Arrest in Two Years	0.13	0.16	0.25	0.72	1.00
Mean	6.16	6.31	0.70	0.21	0.12
Standard Deviation	2.78	2.76	0.46	0.41	0.33
N	24,418	25,009	25,009	25,009	25,009

Note: \*All two-tailed inference tests of  $H_0: \rho=0$ :  $p < .001$ .

Table 7: *Spearman Rank-Order Correlations between Two COMPAS Recidivism Risk Scale Scores (Deciles) and Arrest Type within Two Years of Release (22,153 Male Parolees with Two Year Observation Period)*

	Recidivism Risk Score Decile	Violent Recidivism Risk Score	Any Arrest within Two Years	Arrest for Violent Offense within Two Years	Arrest for Violent Felony within Two Years
Recidivism Risk Score Decile	1.00				
Violent Recidivism Risk Score Decile	0.68	1.00			
Any Arrest in Two Years	0.32	0.28	1.00		
Violent Arrest in Two Years	0.19	0.21	0.34	1.00	
Violent Felony Arrest in Two Years	0.13	0.16	0.25	0.73	1.00
Mean	6.11	6.42	0.71	0.22	0.13
Standard Deviation	2.79	2.74	0.45	0.42	0.34
N	21,187	22,153	22,153	22,153	22,153

Note: \*All two-tailed inference tests of  $H_0: \rho=0$ :  $p < .001$ .

Figure 10: Percentage of Parolees Arrested within Two Years of Parole Release by COMPAS Recidivism Score Decile

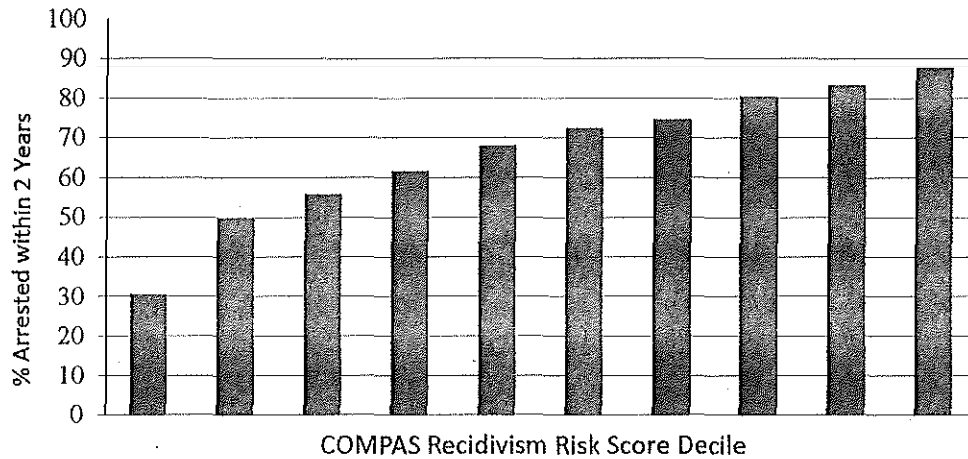
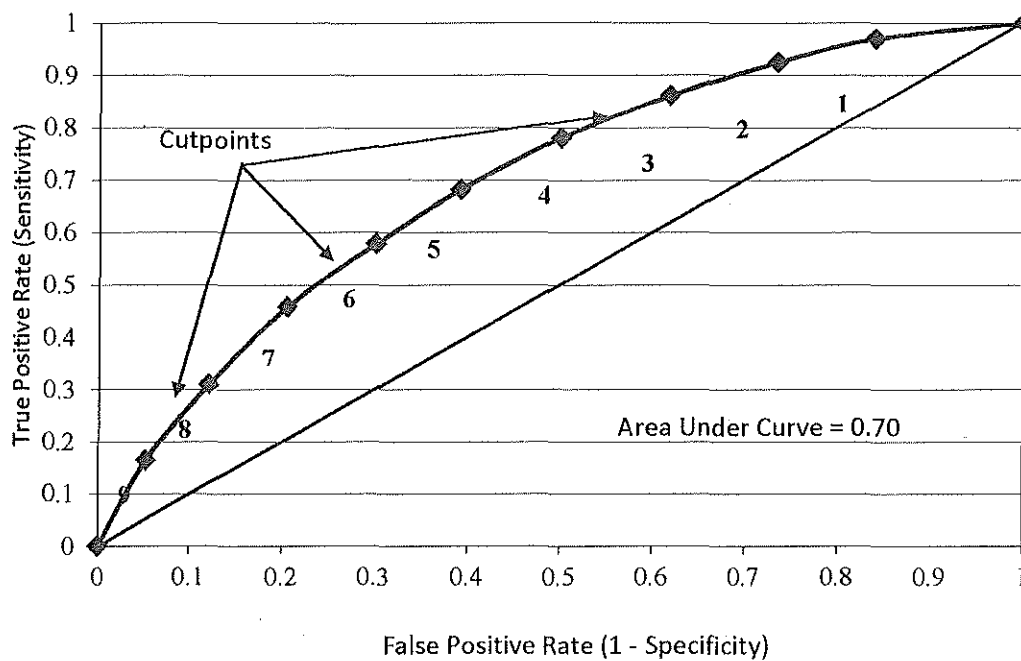


Figure 11: ROC Chart of COMPAS Recidivism Risk Scale Score Decile (Outcome: Arrested within Two Years of Parole Release)



## IV.C.3. Logistic Regression and Comparative Predictive Power Analysis

These analyses examined the predictive power of the COMPAS risk scores in relation to well-known predictors of re-arrest: gender, age, age of first arrest, and the number of prior arrests. To facilitate the analysis, the two-year sample ( $N = 25,009$ ) was randomly divided into two subsamples. One of the samples was used to evaluate changes in the odds of re-arrest with each increment in COMPAS score risk decile. The other sample was used to estimate the predictive power achieved when relying on a small set of background characteristics available electronically from CDCR. Results of these analyses are presented in Tables 9 and 10.

Table 9: Comparison of Odds-Ratios and Test Accuracy from Logistic Regression of Any Arrest within Two Years on COMPAS Recidivism Risk Score Decile Or Select Parolee Characteristics (Parolees with 24 Month Observation Period following Release)

	Model 1	Model 2
Violent Recidivism Risk Score Decile	***1.30	---
Female	---	***0.66
Age at Release	---	***0.93
Age at First Arrest	---	1.00
Total Prior Arrests	---	***1.10
Test Accuracy (c)	0.70	0.72
Likelihood Ratio Chi-Square <sup>1</sup>	***2,531.16	***3,042.95
N	23,805	23,805

Notes: ~:  $p < .10$ ; \*:  $p < .05$ ; \*\*:  $p < .01$ ; \*\*\*:  $p < .001$ ; two-tailed tests.

<sup>1</sup>: Compared to intercept-only model

Table 10: Comparison of Odds-Ratios and Test Accuracy from Logistic Regression of Violent Offense Arrest within Two Years on COMPAS Violent Recidivism Risk Score Decile or Select Parolee Characteristics (Parolees with 24 Month Observation Period following Release)

	Model 1	Model 2
Violent Recidivism Risk Score Decile	***1.24	---
Female	---	***0.55
Age at Release	---	***0.94
Age at First Arrest	---	***0.99
Total Prior Arrests	---	***1.04
Test Accuracy (c)	0.65	0.67
Likelihood Ratio Chi-Square <sup>1</sup>	***1,201.06	***1,454.53
N	25,009	25,009

Notes: ~:  $p < .10$ ; \*:  $p < .05$ ; \*\*:  $p < .01$ ; \*\*\*:  $p < .001$ ; two-tailed tests.

<sup>1</sup>: Compared to intercept-only model

pragmatic question is whether a higher (or lower) threshold will lead to improved programming decisions and ultimately to better supervision outcomes.

Regarding the predictive efficacy of the COMPAS risk scales, our analyses revealed that the COMPAS recidivism and violence scales were significantly correlated with re-arrests during the 24-month follow-up period. However, only the recidivism risk scale achieved the .70 AUC benchmark. Moreover, the risk prediction resulting from the COMPAS scales was comparable to risk prediction models using existing electronic records maintained by CDCR.

The findings in this report should be interpreted in light of two important limitations. The first, as noted above, is that discrepancies between COMPAS- and LSI-R-based needs classifications do not in any way indicate inferiority of the COMPAS. The LSI-R was chosen as a "yardstick" for this evaluation simply because it is a prominent correctional assessment that has been the focus of a number of validity studies. Dramatic deviations from LSI-R classifications might suggest that a new instrument may need further validations with other measurement strategies. However, more subtle discrepancies from the LSI-R are difficult to interpret—and could even indicate superiority of the newer measure(s). The second limitation of this study relates to the outcome variables chosen to validate the risk scales. Re-arrest is an imprecise proxy for post-release criminal activity. Parolees are typically caught and arrested for only a small fraction of crimes/violations they actually commit. As a result, there is a substantial amount of error in the primary outcomes used in these risk-prediction models, which likely reduces their predictive power. However, this problem is partially mitigated by the extremely large parolee sample available for analysis.

Based on these analyses, we conclude that the COMPAS is a reliable instrument with moderate concordance with select LSI-R scales (with significant or marginally significant associations with eight of the nine scales that overlap with the LSI-R). With regard to the predictive validity of the COMPAS risk scales, the general recidivism risk scale achieved the AUC value of .70, which is the conventional standard, though the violence risk scale did not.

Figure 16: Percentage of Parolees that Committed A Violent Parole Violation within Two Years of Parole Release by COMPAS Violent Recidivism Risk Score Decile

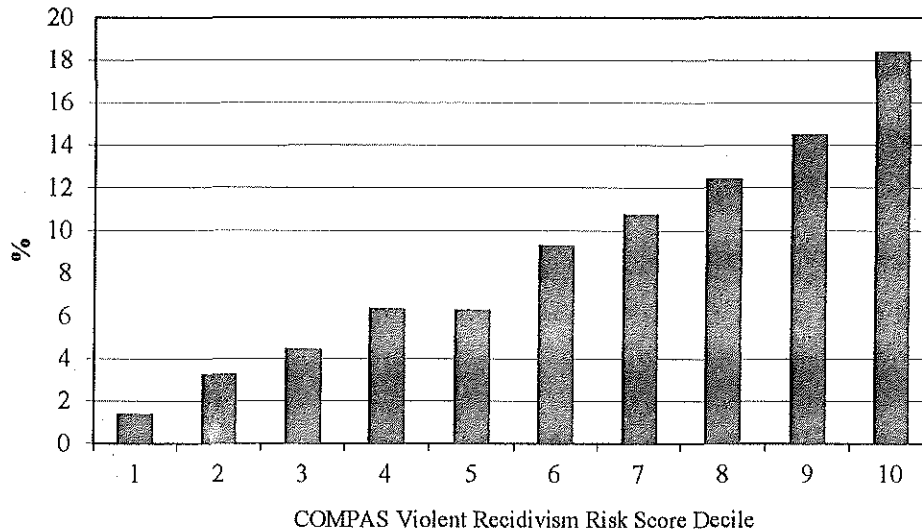


Figure 17: ROC Chart of COMPAS Violent Recidivism Risk Scale Score Decile (Outcome: Violent Parole Violation within Two Years of Parole Release)

