California Pretrial Assessment (CAPA) Validation Study

An outcomes study conducted for the San Diego County Sheriff's Department

Research Department
Northpointe, Inc.

with thanks to William Dieterich

November 16, 2020

equivant
Northpointe Inc. d/b/a equivant

Contents

1 Executive Summary

1.1 Introduction

This report presents validation test results for the California Pretrial Assessment (CAPA) developed by Northpointe in collaboration with the San Diego County Sheriff's Department (SDSO). The Department is responsible for conducting assessments with pretrial defendants booked into the jail system.[1] The results of the pretrial assessments are used by pretrial services and judges to guide release and bail decisions.

CAPA Background

The CAPA was developed in an external sample from a 2010 pretrial study conducted in Kent County Michigan (Dieterich 2010). The development of the CAPA was influenced by California Legislation pertaining to bail reform and the use of pretrial risk assessment tools. Senate Bill 10 (SB-10) and Senate Bill 36 (SB-36) include specifications for pretrial risk assessment tools and pretrial procedure (Senate Bill 10 2018; Senate Bill 36 2019). SB-10 was approved by the Governor and was slated to go into effect on October 1, 2019. Opponents of SB-10 were successful in adding a veto referendum to the November 3, 2020 ballot; the referendum resulted in California voters repealing SB-10. SB-36 was signed into law in November 2019. Details about the the development of the CAPA including the influence of SB-10 and SB-36 is in the Additional CAPA Background Appendix.

Although SB-10 was repealed, both SB-10 and SB-36 include language that reflects ideas advanced by bail reform initiatives around the United States. These ideas have already significantly impacted pretrial practice in California and in particular the development and use of pretrial assessments. Northpointe Research was asked to develop a replacement for the Pretrial Release Risk Scale II (PRRS-II), the Department's current pretrial assessment. The Department required an assessment that was compliant with the pretrial risk assessment tool specifications in SB-10 and SB-36.

---

[1] "Male inmates are booked at either the San Diego Central Jail in downtown San Diego or the Vista Detention Facility in Vista. All female inmates come into the jail system through either the Las Colinas Detention & Reentry Facility in Santee or the Vista Detention Facility in Vista. Field bookings are not normally accepted at any other jail." (San Diego Sheriff's Department 2019)

Development and Validation Steps

The following steps were taken to develop and validate the CAPA:

- Developed the CAPA in the training data from the 2010 pretrial study conducted in Kent County Michigan following SB-10 specifications.

- Added five test items to scale.

- Added the CAPA and test items to the SDSO COMPAS application.

- Collected pretrial assessment data with the CAPA including test items on a representative SDSO sample of 7,987 defendants released pretrial.

- Updated the CAPA by dropping one item, replacing one item with a variant test item, and adding another test item.

- Tested the discriminative and predictive ability of the CAPA in the SDSO sample.

- Examined failure to appear (FTA) and new criminal activity (NCA) outcomes separately as well as an aggregate outcome.

- Examined race and gender effects.

- Set cuts in the SDSO validation sample to form Low, Medium, and High Levels as specified in SB-36.

1.2 Study Overview

Standard measures of predictive validity were applied to evaluate the discriminative ability and predictive ability of the CAPA. Discriminative ability was assessed with the area under the receiver operating characteristic (ROC) curve (AUC). Predictive ability was assessed by evaluating the positive predictive values at scale cut points. The validation was conducted in a sample of 3,953 CAPA assessed individuals who were released pretrial. The discriminative and predictive validity of the risk scales was tested in gender and ethnic groups.

1.3 Study Goals and Objectives

The goals of the study are (1) to test the validity of the CAPA and (2) to test for differential validity of the CAPA in gender and ethnic groups.

The study has the following objectives:

1. Test the ability of the CAPA to discriminate new criminal arrests (NCA) and failure to appear for a scheduled court hearing (FTA), separately and in combination.[2]

2. Evaluate the predictive ability of the CAPA at the High cut-point.

3. Compare discriminative ability and predictive ability for gender and ethnic groups.

---

[2]NCA is typically used in the pretrial literature to designate new criminal activity. We borrow the same abbreviation, but in our study NCA designates new criminal arrest as opposed to new criminal activity. In some pretrial studies new arrest (NA) is used instead of NCA.

1.4   Major Findings

Validity of the CAPA

Discriminative Ability

The ability of the CAPA score to discriminate NCA, FTA, and the aggregate outcome NCA/FTA during pretrial release was tested. Discriminative ability was measured with the area under the receiver operating characteristic (ROC) curve (AUC). The AUC is a summary measure of discrimination across all thresholds of a scale. Defined loosely, AUC is the probability that a randomly selected person who goes on to recidivate will have a higher risk score than a randomly selected person who does not go on to recidivate. AUC can take on values between 0 and 1, with values around 0.5 indicating that the scale discriminates no better than a coin toss. According to the guidelines provided by Desmarais and Singh (2013), AUC values of .50 to .54 are considered poor, .55 to .63 fair, .64 to .70 good, and .71 and higher are excellent. More conventional standards used in criminal justice applications state that AUCs of .65 to .69 indicate acceptable discrimination and AUCs in the range .70 to .75 indicate good discrimination. Additional technical details are in the methods section.

Discriminative Ability for NCA. The AUC obtained for the CAPA in the overall sample is 0.699 (95% CI: 0.675-0.724). The AUC is 0.706 (95% CI: 0.659-0.754) for women and 0.696 (95% CI: 0.667-0.725) for men. Thus, the CAPA has good discriminative ability in the sample overall and in the gender groups. Furthermore, the AUCs for women and men are not statistically different, nor are the AUCs across ethnic groups. More details are shown in Table 8.

Discriminative Ability for FTA. The AUC obtained for the CAPA in the overall sample is 0.643 (95% CI: 0.617-0.669). The AUC is 0.700 (95% CI: 0.653-0.748) for women and 0.621 (95% CI: 0.590-0.651) for men. Thus, the CAPA has fair discriminative ability in the sample overall and for men, with good discriminative ability for women. The AUCs for women and men are statistically different ($p = 0.006$), but they are not different across ethnic groups. More details are shown in Table 9.

Discriminative Ability for aggregate NCA/FTA. The AUC obtained for the CAPA in the overall sample is 0.685 (95% CI: 0.666-0.705). The AUC is 0.726 (95% CI: 0.691-0.762) for women and 0.669 (95% CI: 0.645-0.692) for men. Thus, the CAPA has fair discriminative ability in the sample overall and for men, with good discriminative ability for women. The AUCs for women and men are statistically different ($p = 0.008$), but they are not different across ethnic groups. More details are shown in Table 10.

Predictive Ability

The results demonstrate that the CAPA Risk Levels separate the pretrial study sample into groups with distinct failure rates. The overall probability of NCA given a High Risk score (scores of 6 and higher) is approximately 17 percentage points higher than the base rate of 12%, the overall probability of FTA given a High Risk score is approximately 14 percentage points higher than the base rate of 13%, and the overall probability of either NCA or FTA given a High Risk score is approximately 25 points higher than the base rate of 23%. This

indicates that the CAPA Risk Levels have good predictive ability and are clinically useful.

Predictive Validity for Genders and Ethnic Groups

During follow up, similar proportions of women (29%) and men (29%) in the High Risk Level experience an NCA; 30% of women and 26% of men in the High Risk Level experience an FTA; 53% of women and 50% of men in the High Risk Level experience either an NCA or FTA. These figures demonstrate predictive parity and fairness in the High Risk Level across genders for the various outcomes. Likewise, similar proportions across ethnicities in the High Risk Level fail (NCA or FTA or combined) during the follow-up, demonstrating predictive parity and fairness in the High Risk Level across ethnicity for the three outcomes considered. Details are discussed in Section 3.3.

The findings in the outcomes study demonstrate the predictive validity of the CAPA risk scales in the SDSO application and provide strong support for their use by the Department to guide supervision level and conditions of pre-trial release.

2 Materials and Methods

2.1 California Pretrial Assessment

A pilot version of the CAPA was developed prior to the start of the present study in a training data set from a 2010 pretrial outcomes study conducted in Kent County Michigan. The pretrial assessment included the seven items needed to score the CAPA plus five test items that were not scored. The pilot version of the CAPA was updated by dropping cc_any_arrest_on_bail_pts (Has the person been arrested/charged with a new crime that resulted in a conviction while on pretrial release?), replacing cc_n_fta_pts (How many times has this person failed to appear for a scheduled criminal court hearing?) with the test item cc_n_fta_2_pts (How many times has this person failed to appear for a scheduled criminal court hearing within the last two years?), and adding the test item cc_n_pconviction_pts (How many times has this person been convicted for a misdemeanor or felony before as an adult?). The updated CAPA instrument that we test in the current study is in CAPA Items Appendix.

2.2 Sample

SDSO pretrial services conducted CAPA assessments on consecutive bookings into the SDSO jail system. A total of 7,987 persons with a pretrial jail booking were assessed with the CAPA. The CAPA assessments were conducted between September 30, 2018 and December 12, 2018. The target for the study was a sample of 5,000 assessed persons released pretrial. In the assessed sample, 2,450 persons had their cases closed, 1,009 persons were detained, 157 persons were missing one or more pretrial data elements, and 4,371 persons were released pretrial. After dropping persons with fewer than 14 days on pretrial release, persons arrested for an offense committed prior to the index pretrial release, and persons older than 79, the sample included 3,953 defendants released pretrial. We thank the IT department at SDSO for assisting with identifying the bookings in the data that could not be followed reliably.

The main sample for analysis consists of the 3,953 individuals with an intact CAPA assessment that were released pretrial. Men represent 72.1% of the main sample (n=2,849), and women represent 27.9% of the main sample (n=1,104). The median age at assessment is 33.1 ($Mean = 35.9$). The median age at first arrest is 25 ($Mean = 28.6$). The mean number of

prior misdemeanor convictions is 1.4. The racial/ethnic composition of the sample is 44% Caucasian, 13% Black, 36% Latino, and 7.1% other ethnic/racial groups.

Release dates for the sample ranged from September 30, 2018 to December 12, 2018. The end of study was set to June 28, 2019. The median number of days between release and either failure, disposition, or end of study was 140 days.

Table 1 shows the counts and percentages of pretrial index arrest top charge by charge class in the full sample ($n = 7,981$). The index arrest top charge category with the highest frequency is drug offense (29.3%), followed by alcohol (16.0%), assault (15.1%), property (14.8%), and domestic violence (7.0%).

2.3 Study Outcomes

Three types of pretrial failure are examined: 1) an arrest for a misdemeanor or felony offense that occurred during the pretrial episode, 2) failure to appear for a scheduled court hearing (FTA) during the pretrial episode, and 3) the aggregate NCA/FTA. The failure time point is the offense date for NCA and warrant date for FTA.

SDSO Data Services obtained the study outcomes for failure to appear (FTA); new arrest on pretrial (NCA); and disposition from the Jail Information Management System (JIMS), the District Attorney database (DA), and the criminal case management system (JURIS). Matching on booking number and arrest date, we were able to match all released persons to an index case record in at least one of the three databases (JIMS, DA, or JURIS).

The pretrial release episode ends on the date of FTA bench warrant, date of bail bond forfeiture, date of new felony arrest, date of new misdemeanor arrest, new jail booking for violations of bond conditions, index case disposition date, or the end of study, whichever occurs first.

A pretrial failure is defined only for charges or offenses that occurred during the pretrial follow-up period. Pending warrants and charges for anyone entering the study will not constitute a failure if executed during the pretrial follow-up period.

Defendants that are released pretrial but have their case dismissed (without pretrial misconduct) will be counted as a success if they accumulate more than 14 days of pretrial follow-up. Cases that don't have at least 14 days of pretrial follow-up (meaning they probably had their case dismissed at the preliminary examination) will not be included in the study. Defendants that never bond out but have their case dismissed at the preliminary examination will not be included in the study, either.

Released cases for whom no charges were filed by the arraignment date were not considered as pretrial releases and were excluded from the outcomes analysis.

2.4 Analytical Approach

The ability of the CAPA score to discriminate NCA and FTA during pretrial release was tested. Discriminative ability was measured with the area under the receiver operating characteristic

(ROC) curve (AUC). The AUC is a summary measure of discrimination across all thresholds of a scale. By convention, in criminal justice applications, AUCs of .65 to .69 indicate acceptable discrimination and AUCs in the range .70 to .75 indicate good discrimination. Discriminative ability is related to diagnostic accuracy. Measures of discriminative ability are widely used in the medical field to evaluate diagnostic tests and scales. Diagnostic accuracy is often confused with predictive accuracy. The AUC is a rank-based measure of discriminative ability, not a measure of accuracy (e.g. Steyerberg et al. 2010). There has been an over-reliance on the AUC as a measure of risk scale performance in criminal justice and other fields. The shortcomings of the AUC have been noted and discussed widely (Guggenmoos-Holzmann and Houwelingen 2000; Singh 2013; Levy 2020; Royal Statistical Society Section on Statistics and the Law 2018). We rely on the AUC as a measure of discrimination and at the same time try to demote it from its elevated status and guard against misinterpreting it as a measure of predictive accuracy.

The predictive ability of the CAPA was evaluated with the positive predictive values. The positive predictive value (PPV) is the probability that a person with high risk score will fail. The negative predictive value (NPV) is the probability that a person without a high risk score will not recidivate. PPV and NPV assess predictive ability. A useful prediction will have a PPV that is greater than the base rate (overall failure rate) and a NPV that is greater than 1 minus the base rate. A perfect test will predict the outcome perfectly with PPV = 1 and NPV = 1. The predictive values depend on the accuracy of the test and the base rate.

Table 1: Index pretrial arrest offense group.

	Offense Class			
	Infraction	Misdemeanor	Felony	All
Homicide	0	0	26	26
	0.00	0.00	1.03	0.33
Sex Offense	0	26	52	78
	0.00	0.48	2.06	0.98
Violent Property	0	0	265	265
	0.00	0.00	10.52	3.32
Assault	0	782	424	1,206
	0.00	14.51	16.83	15.11
Domestic Violence	0	40	522	562
	0.00	0.74	20.71	7.04
Weapon Offense	0	86	210	296
	0.00	1.60	8.33	3.71
Property	25	640	519	1,184
	34.72	11.88	20.60	14.84
Drug Offense	0	2,018	323	2,341
	0.00	37.45	12.82	29.33
Escape	0	0	7	7
	0.00	0.00	0.28	0.09
Other Domestic Viol.	0	193	0	193
	0.00	3.58	0.00	2.42
Alcohol	1	1,222	55	1,278
	1.39	22.68	2.18	16.01
Other Offense	46	382	117	545
	63.89	7.09	4.64	6.83
Total	72	5,389	2,520	7,981
	100.00	100.00	100.00	100.00

3   Results

3.1   Comparison of Released, Detained, and Closed Cases

As is the case in all pretrial release samples, the SDSO pretrial study sample is affected by selection mechanisms that determine which defendants are released and included in the estimation sample. To be included in the estimation sample, defendants must be assessed with CAPA and released pretrial (bail, bond, ROR, or SPTS). The CAPA assessments were conducted with consecutive pretrial bookings which minimizes selection effects at the assessment stage. However, subsequent release and bail decisions determined the composition of the estimation sample. Released defendants differ in important ways from detained defendants. CAPA was not used to guide pretrial release decisions in the present study, but some of the factors considered by judges and other stakeholders in making their bail and release decisions are similar to those included in CAPA. Judges and stakeholders also rely on other information not covered by CAPA, much of which is unknown to us.

We have CAPA scores for the full sample, but we only observe the pretrial outcomes of defendants that were released. The pretrial outcomes for detained persons are unobserved. In the present study, we will be testing the predictive and discriminative ability of the CAPA score in the group of defendants that were released pretrial. We will not determine the counterfactual - how the CAPA would perform if the release sample also included the detained defendants. There are special selection models that are sometimes useful for inferential statistics, but those models will not be helpful in the present validation work. It is however informative to explore the selection effects.

Table 2 shows the counts and percentages of release status by the pretrial index arrest top charge classification. Detained cases have a much higher proportion of felony charges than the other categories of pretrial release status, while most cases with an infraction for top charge class are closed.

Table 2: Pretrial release status by index offense class.

	Pretrial Release Status				
	Released	Detained	Closed	Missing	All
Infraction	1	1	68	2	72
	0.02	0.10	2.77	2.60	0.90
Misdemeanor	3,212	247	1,869	61	5,389
	72.31	24.50	76.16	79.22	67.52
Felony	1,229	760	517	14	2,520
	27.67	75.40	21.07	18.18	31.57
Total	4,442	1,008	2,454	77	7,981
	100.00	100.00	100.00	100.00	100.00

Selection effects were explored by comparing the released defendants, defendants that were not released, and defendants that had their cases closed before the first arraignment. Figure 1 compares these three groups with respect to several criminal history and demographic variables (e.g. race; fta; priors). The study variables are either naturally dichotomous (e.g. Male=0, Male=1) or they were changed to dichotomous variables (e.g. n_fta_pts: number of previous FTAs. This variable was originally recorded as 0 times, 1 time, or at least 2 times. The new dichotomous variable records this variable as 0 times or at least 1 time). The variables are sorted in descending order of variable importance (how well they discriminate between release groups). The plot points have error bars showing the 95% confidence interval (CI) around the point estimate of the proportion. The top six factors that discriminate detained from released defendants are current felony, prior conviction, supervision status, prior FTA, pending charge, and prior jail. In comparison with detained defendants, defendants that had their cases closed were less likely to have a current felony, current larceny, or to be male and more likely to be Anglo. There are other statistically significant differences but only a few of these are practically significant (large enough to warrant action).

Table 3 shows the counts and percentages of release status by pretrial index arrest top charge class for men and women. A somewhat higher percentage of men than women were released on recognizance (ROR) ($p = 0.0054$).
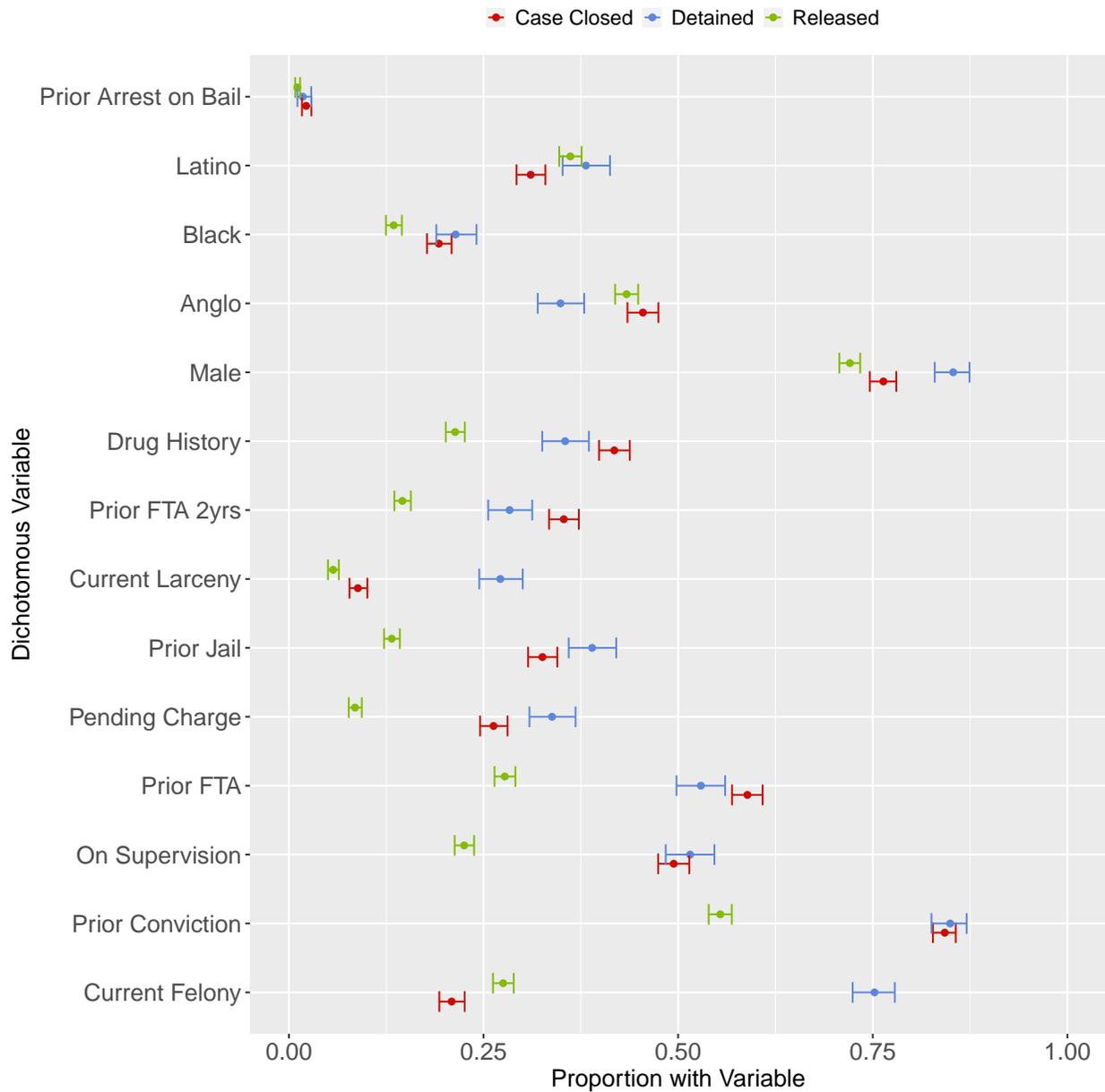
Figure 1: Comparison of study and demographic factors for defendants released pretrial with defendants that were detained or had their cases closed.

Table 3: Counts and percentages of release types for women and men in the pretrial release sample.

	Gender		
	Women	Men	Sum
B & R	131	314	445
	11.87	11.02	11.26
Bond	836	2,112	2,948
	75.72	74.13	74.58
Cash	25	49	74
	2.26	1.72	1.87
ROR	85	305	390
	7.70	10.71	9.87
SPTS	27	69	96
	2.45	2.42	2.43
Total	1,104	2,849	3,953
	100.00	100.00	100.00

Table 4 shows the counts and percentages of release status by pretrial index arrest top charge class for Anglo, Black, and Latino defendants. A chi-squared test indicates that release type is not independent across ethnicities. Proportions of Blacks and Anglos released on bond are lower than expected, while the proportion of Latinos released on bond is higher. A higher than expected proportion of Anglos are released with B & R while a higher than expected proportion of Blacks are released on recognizance. By contrast, a lower than expected proportion of Latinos are released with either of these two methods.

Table 4: Counts and percentages of release types for ethnic groups in the pretrial release sample.

	Ethnic Group				
	Anglo	Black	Latino	Other	Sum
B & R	244	62	123	16	445
	14.04	12.11	8.65	5.69	11.26
Bond	1,265	356	1,104	223	2,948
	72.78	69.53	77.64	79.36	74.58
Cash	27	3	32	12	74
	1.55	0.59	2.25	4.27	1.87
ROR	172	77	117	24	390
	9.90	15.04	8.23	8.54	9.87
SPTS	30	14	46	6	96
	1.73	2.73	3.23	2.14	2.43
Total	1,738	512	1,422	281	3,953
	100.00	100.00	100.00	100.00	100.00

3.2   Proportions in the CAPA Raw Scores and Levels

Figure 2 displays bar plots with the percentage of women and men in the CAPA raw scores. The bar plots show that the number of men within scores generally decreases as the score increases. The score distribution of women is similar, with a somewhat larger proportion of women than men obtaining the lowest possible score.
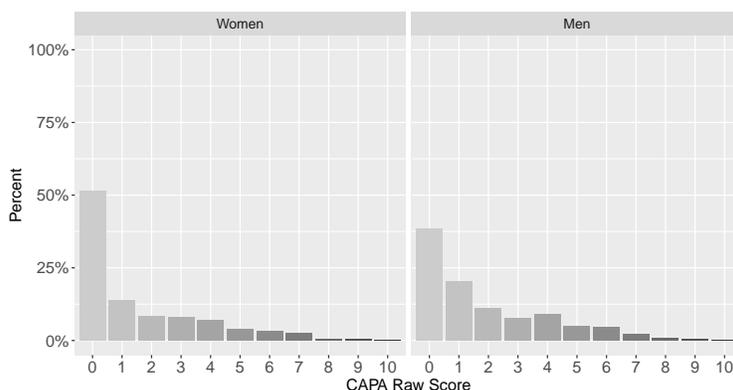
Figure 2: Percentage of women and men in the CAPA scores ($n =$3,953).

Figure 3 shows the percentage defendants in the CAPA raw scores across ethnic groups. The bar plots show that the proportion of individuals within scores generally decreases as the score increases, although there is an unexpected density at scores of 4. The score distribution across ethnicities are similar.
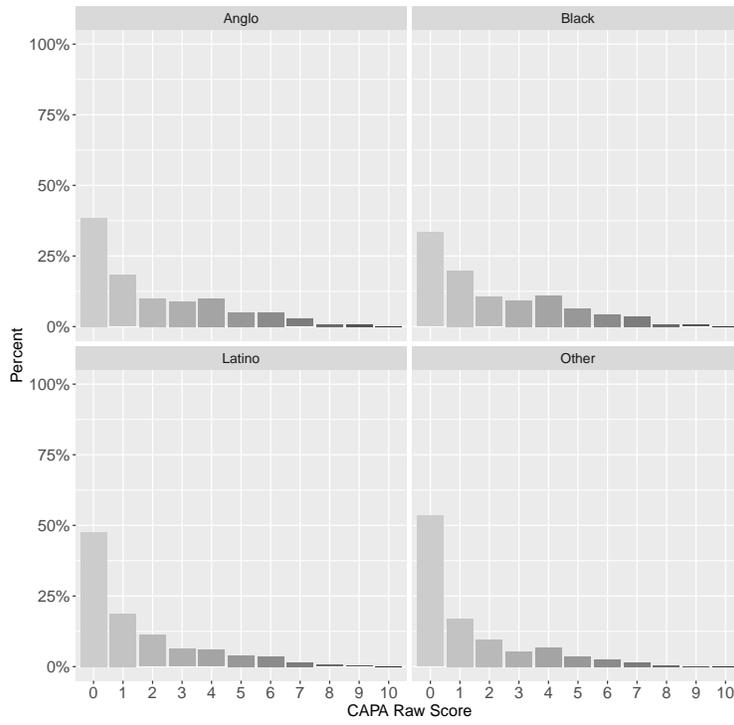
Figure 3: Percentage of defendants in the CAPA scores by ethnic group ($n =$ 3,953).
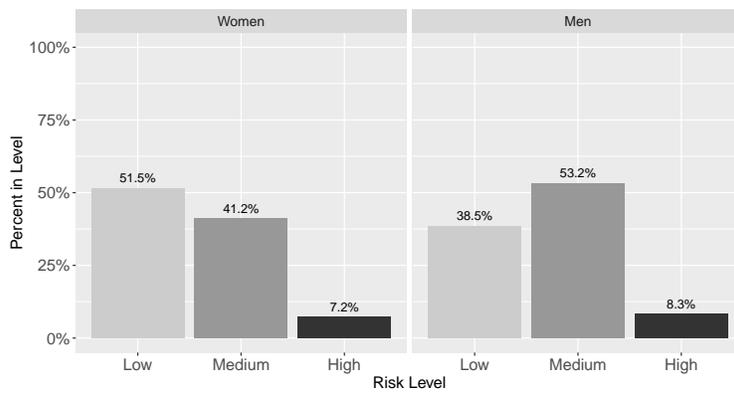
Figure 4: Percentage of women and men in the levels of the CAPA.
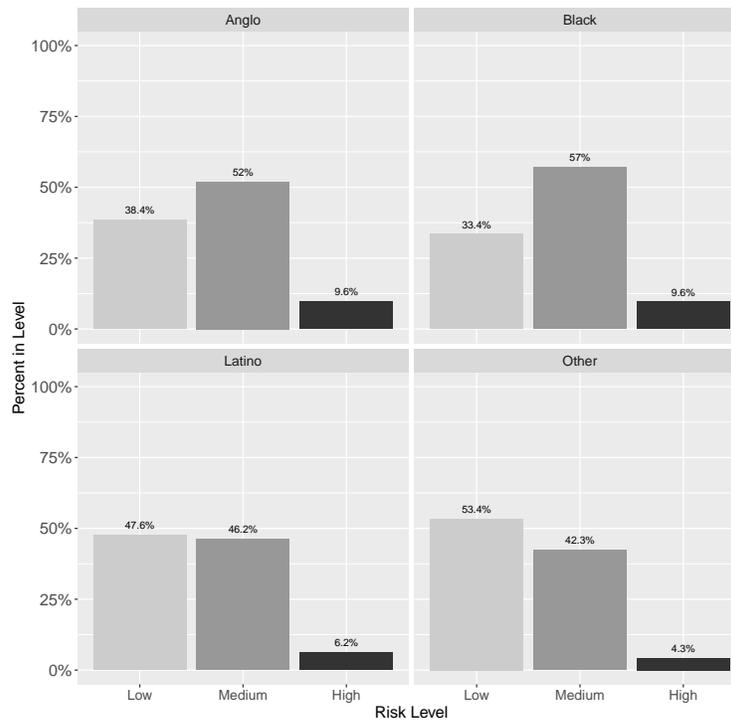
Figure 5: Percentage of Anglo, Black, Latino, and Other ethnic group defendants in the levels of the CAPA.

The risk level for individuals obtaining CAPA scores of 0 was set at Low, the risk level for individuals scoring between 1 and 5 was set at Medium, and the risk level for those scoring 6 to 11 on the CAPA was set at high. The distribution of risk levels for men and women is shown in Figure 4. Similar proportions of women (7.2%) and men (8.3%) are in the High risk group, while a somewhat higher proportion of men than women are Medium risk.

Figure 5 displays the distributions of risk levels across ethnicities. These distributions are similar in that the High risk level contains the smallest proportion of individuals, regardless of ethnicity, with about 9.6% of Anglos and Blacks considered High Risk, and even smaller proportions for Latinos (6.2%) and Others (4.3%). About 57% of Blacks are considered Medium risk; a somewhat smaller proportion of Anglos (52%) are considered Medium risk, and only 46.2% of Latinos are Medium risk. 53.4% of those whose ethnicity is classified as "Other" were assessed as Low risk. 47.6% of Latinos were assessed as Low risk, while 38.4% of Anglos and 33.4% of Blacks were in this risk level.

Table 5: Counts and percentages of events observed during pretrial release.

| | *Index Offense Class* | | |
	Misdemeanor	Felony	Sum
Censored	1,243	580	1,823
	42.89	55.03	46.12
New Arrest	395	85	481
	13.63	8.06	12.17
FTA	307	138	445
	10.59	13.09	11.26
Disposition	953	251	1,204
	32.88	23.81	30.46
Total	2,898	1,054	3,953
	100.00	100.00	100.00

3.3 Pretrial Release Outcomes

The counts and percentages of the events observed during pretrial release by pretrial index offense class are in Table 5. "Censored" means that by the end of the study (June 28, 2019), these cases were still not disposed and no NCA or FTA had occurred.

Pretrial outcomes are more complex than simple arrest recidivism outcomes. The event of interest might be NCA, but there are other events that compete with the event of interest and alter the probability of observing it. NCA and FTA can only occur during the pretrial release period. The pretrial release period ends when the index case is disposed, the pretrial release is revoked, or the study ends. Case disposition is a competing event that alters the probability of observing NCA or FTA. An NCA that leads to incarceration will alter the probability of observing an FTA. An FTA that that leads to a bond revocation and detention will compete with NCA. A special type of survival analysis can be used to obtain estimates of failure probabilities when competing events are present. Simply ignoring competing events can lead to biased probability estimates. Competing events can also affect measures of discrimination such as the AUC. Generally when analyzing competing risks the cumulative incidence function should be used to estimate failure probabilities as opposed to the simple proportion failing or the naive 1-Kaplan-Meier failure probabilities. The cumulative incidence is the probability of failing by the event of interest adjusted by the overall survival probability for all events.

Survival models might be appropriate for the pretrial study data because we are interested in both the occurrence and timing of the pretrial outcomes. Prior research has shown that the risk of pretrial misconduct increases as the time period from pretrial release to final case disposition increases (Cohen and Reaves 2004). The method is well suited to modeling the hazard of pretrial arrest and competing failure events in the pretrial setting. But the survival model is problematic for FTA because defendants are only at risk of FTA when

they have a scheduled court appearance (Visher and Linster 1990; Rhodes, Hyatt, and Scheiman 1996). Rhodes, Hyatt, and Scheiman (1996) discussed several methodological and statistical challenges when conducting pretrial outcomes analysis. We list a few of the salient methodological challenges here: 1) FTAs are only possible on scheduled hearing dates, 2) the probability of pretrial misconduct is conditioned by the number of hearings, 3) the probability of FTA may be higher for potentially hazardous court events such as sentencing hearings, and 4) the probability of pretrial misconduct is conditioned by the number of days from pretrial release to disposition or end of study.

Competing risk analysis that takes competing risks into account may be more important for model development than model validation. We know of no satisfactory approach that has been demonstrated for fitting survival models to events such as FTA. Our general approach for this study is to treat NCA and FTA as binary outcomes (success or failure) as opposed to modeling failure time.

We consider three possible pretrial release outcomes: NCA, FTA, or case disposition.

- Events of interest: NCA, FTA, or aggregate outcome NCA/FTA.
- Competing events: Successful completion of pretrial release (case disposition).
- Censoring: Case not yet disposed and still pending by the study end date (June 28, 2019).
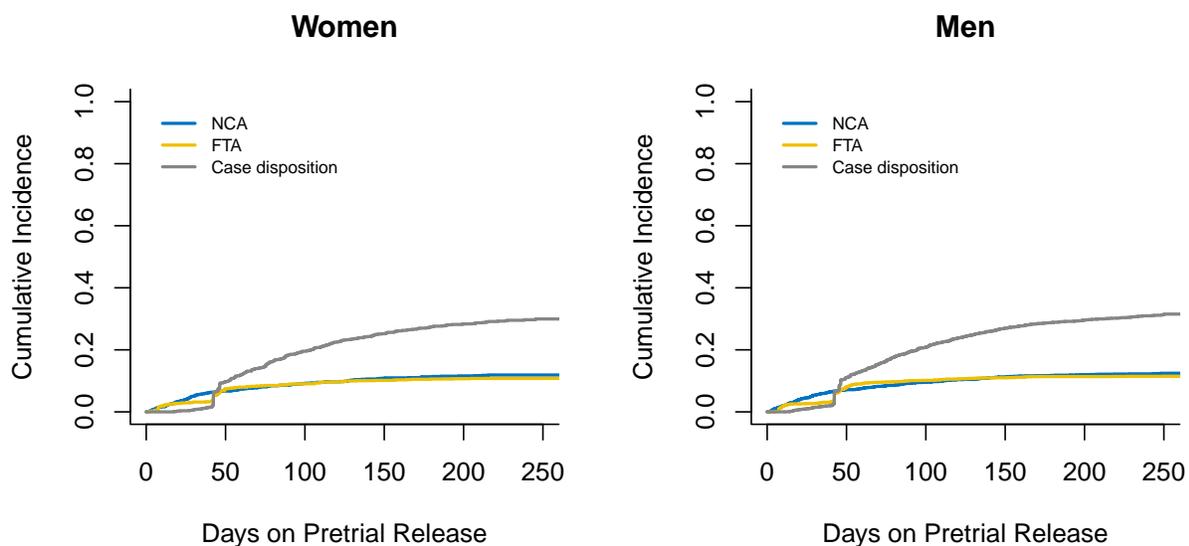
Figure 6: Cumulative incidence curves for new arrest, failure to appear, or case disposition in the SDSO pretrial study sample by gender, excluding censored cases.

Figure 6 shows the cumulative incidence curves for men and women in the pretrial release sample, and Figure 7 shows plots of the crude cumulative incidence curves of NCA in the presence of FTA and case disposition for Anglo, Latino, and Black defendants. The step in the curve for disposition reflects defendants reaching case disposition before FTA or NCA. Here case disposition and FTA alter the probability of arrest on pretrial. The cumulative

incidence is the probability of experiencing an event over time. Overall, the cumulative incidence of pre-trial NCA or FTA by 50 days is 7.2%, lower than the cumulative incidence of case disposal.
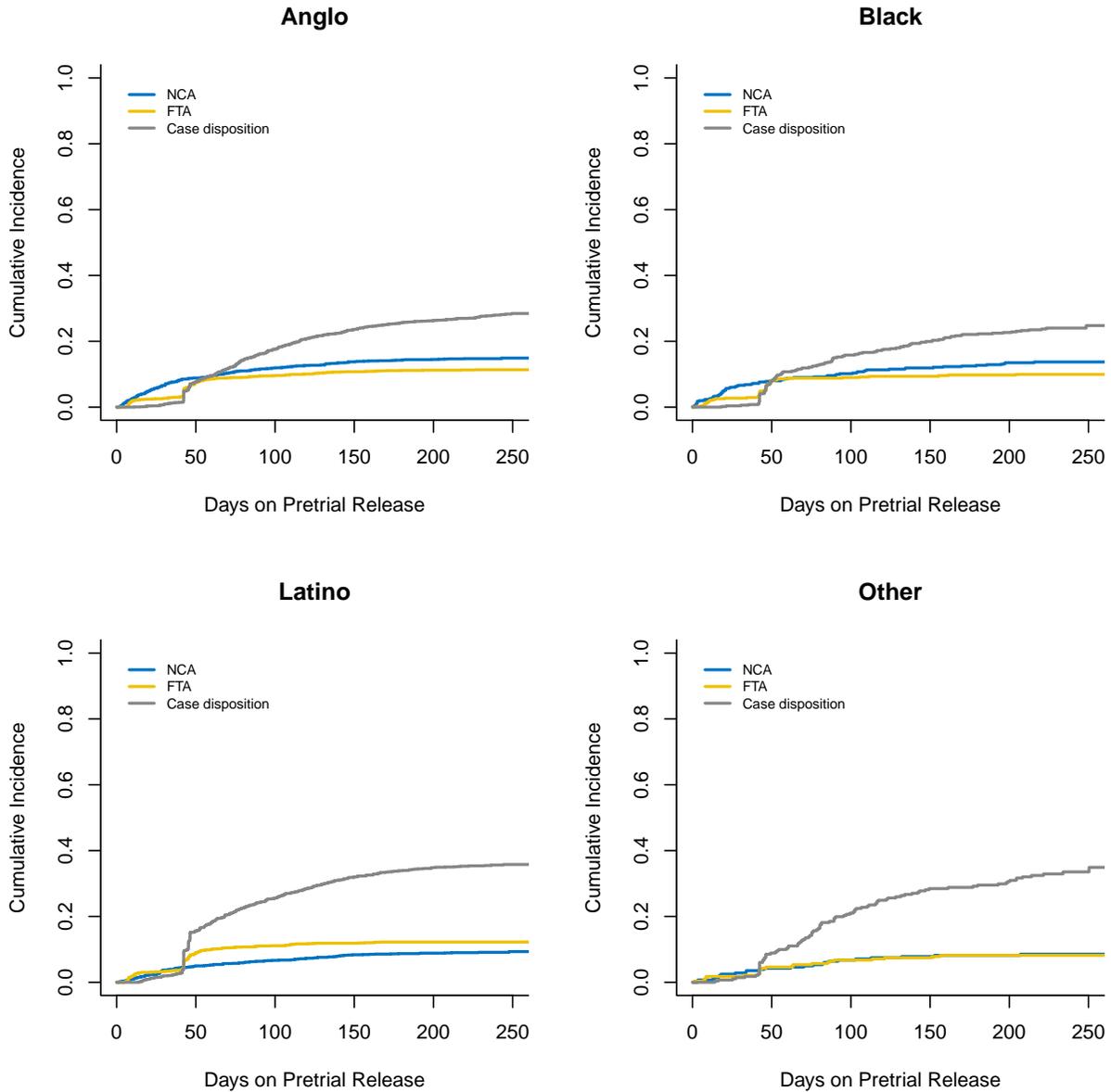
Figure 7: Cumulative incidence curves for new arrest, failure to appear, or case disposition in the SDSO pretrial study sample by ethnic group, excluding censored cases.

Table 6: Descriptive and bivariate statistics for pretrial study factors by ethnic group.

| | | *Anglo* | | | | *Latino* | | | | *Black* | | | |
| | | Total | | NCA/FTA | | Total | | NCA/FTA | | Total | | NCA/FTA | |
		N	%	N	%	N	%	N	%	N	%	N	%
Pending charge	Yes	139	8.0	72	51.8	97	6.8	36	37.1	51	10.0	22	43.1
	No	1599	92.0	382	23.9	1325	93.2	268	20.2	461	90.0	99	21.5
Larceny	Yes	76	4.4	32	42.1	78	5.5	26	33.3	32	6.2	13	40.6
	No	1662	95.6	422	25.4	1344	94.5	278	20.7	480	93.8	108	22.5
Prior jail	Yes	250	14.4	106	42.4	128	9.0	44	34.4	92	18.0	26	28.3
	No	1488	85.6	348	23.4	1294	91.0	260	20.1	420	82.0	95	22.6
Prior conviction	Yes	1013	58.3	335	33.1	680	47.8	188	27.6	319	62.3	84	26.3
	No	725	41.7	119	16.4	742	52.2	116	15.6	193	37.7	37	19.2
Drug history	Yes	433	24.9	191	44.1	211	14.8	82	38.9	126	24.6	41	32.5
	No	1305	75.1	263	20.2	1211	85.2	222	18.3	386	75.4	80	20.7
On supervision	Yes	434	25.0	189	43.5	262	18.4	99	37.8	124	24.2	46	37.1
	No	1304	75.0	265	20.3	1160	81.6	205	17.7	388	75.8	75	19.3
Arrest on bail	Yes	18	1.0	16	88.9	13	0.9	6	46.2	5	1.0	1	20.0
	No	1720	99.0	438	25.5	1409	99.1	298	21.1	507	99.0	120	23.7
FTA total	Yes	486	28.0	200	41.2	286	20.1	112	39.2	182	35.5	58	31.9
	No	1252	72.0	254	20.3	1136	79.9	192	16.9	330	64.5	63	19.1
FTA w/i 2 yrs.	Yes	253	14.6	135	53.4	147	10.3	72	49.0	81	15.8	35	43.2
	No	1485	85.4	319	21.5	1275	89.7	232	18.2	431	84.2	86	20.0
Felony	Yes	426	24.5	99	23.2	391	27.5	81	20.7	131	25.6	26	19.8
	No	1312	75.5	355	27.1	1031	72.5	223	21.6	381	74.4	95	24.9

Table 7: Descriptive and bivariate statistics for pretrial study factors by gender.

| | | *Women* | | | | *Men* | | | |
| | | Total | | NCA/FTA | | Total | | NCA/FTA | |
		N	%	N	%	N	%	N	%
Pending charge	Yes	69	6.2	39	56.5	231	8.1	98	42.4
	No	1035	93.8	211	20.4	2618	91.9	578	22.1
Larceny	Yes	46	4.2	19	41.3	150	5.3	55	36.7
	No	1058	95.8	231	21.8	2699	94.7	621	23.0
Prior jail	Yes	101	9.1	45	44.6	399	14.0	142	35.6
	No	1003	90.9	205	20.4	2450	86.0	534	21.8
Prior conviction	Yes	489	44.3	166	33.9	1645	57.7	478	29.1
	No	615	55.7	84	13.7	1204	42.3	198	16.4
Drug history	Yes	205	18.6	95	46.3	605	21.2	234	38.7
	No	899	81.4	155	17.2	2244	78.8	442	19.7
On supervision	Yes	216	19.6	91	42.1	646	22.7	258	39.9
	No	888	80.4	159	17.9	2203	77.3	418	19.0
Arrest on bail	Yes	9	0.8	6	66.7	29	1.0	19	65.5
	No	1095	99.2	244	22.3	2820	99.0	657	23.3
FTA total	Yes	261	23.6	115	44.1	744	26.1	272	36.6
	No	843	76.4	135	16.0	2105	73.9	404	19.2
FTA w/i 2 yrs.	Yes	143	13.0	77	53.8	361	12.7	178	49.3
	No	961	87.0	173	18.0	2488	87.3	498	20.0
Felony	Yes	277	25.1	56	20.2	777	27.3	167	21.5
	No	827	74.9	194	23.5	2072	72.7	509	24.6

3.3.1 Pretrial study factors

Table 6 compares the pretrial study factors for Anglo, Latino, and Black defendants in the release sample and Table 7 compares the pretrial study factors for male and female defendants in the release sample. The table columns labeled "%" under "NCA/FTA" for each subgroup gives an indication of the association of each factor with pretrial failure. For example, in Table 6, 51.8% of Anglos who had a pending charge at the time of arrest eventually failed with either an NCA or FTA by the time the study ended, compared to only 23.9% of Anglos who did not have a pending charge at the time of arrest. There is a similarly higher proportion who endorse this factor across ethnic groups (and gender- see Table 7) among those who eventually failed compared to those who do not fail within the study period.

The factors "FTA total" (person had any previous FTA) and "FTA within 2 years" (person had a previous FTA within the last two years) both have higher proportions of endorsements among those who fail within the study period, but the differences are greater for "FTA within 2 years." This provides the justification for using this variable in the CAPA rather than "FTA total."

The factor "Arrest on bail" has a lower proportion of endorsements among Black defendants who fail within the study period compared to Black defendants who do not fail. Including this variable among the CAPA factors would have led to incorrectly increased risk scores among Black defendants, so it is not included in the set of CAPA pretrial risk factors.

3.3.2   NCA Outcomes

Discriminative Ability for NCA. The AUC obtained for the CAPA in the overall sample is 0.699 (95% CI: 0.675-0.724). The AUC is 0.706 (95% CI: 0.659-0.754) for women and 0.696 (95% CI: 0.667-0.725) for men. Thus, the CAPA has fair to good discriminative ability in the sample overall and in the gender groups. Furthermore, the AUCs for women and men are not statistically different, nor are the AUCs across ethnic groups. Table 8 shows these figures for all groups, as well as the number in each subgroup, the number with NCA in each subgroup, and the percent with NCA in each subgroup.

Table 8: CAPA AUCs for NCA outcomes by overall sample and subgroups.

	AUC	Low	High	Number failing	N	Percent failing
Overall	0.699	0.675	0.724	481	3953	12.2
Women	0.706	0.659	0.754	131	1104	11.9
Men	0.696	0.667	0.725	350	2849	12.3
Anglo	0.711	0.677	0.745	257	1738	14.8
Black	0.660	0.593	0.726	70	512	13.7
Latino	0.663	0.614	0.712	130	1422	9.1
Other	0.766	0.673	0.858	24	281	8.5

Predictive Ability for NCA. The predicted probability of NCA given CAPA score for men and women is shown in Figure 8. The gray region around each line is the 95% confidence band. Since these confidence bands overlap across all CAPA scores (except 11, since no one in the sample obtained the highest possible score), we conclude that the tool predicts NCA equally well for men and women across CAPA scores.

In general, the failure rate for defendants in the High level is known as the Positive Predictive Value (PPV). Figure 9 shows NCA rates for released persons by gender given their CAPA levels. Error bars are added to indicate the respective confidence intervals. The base rate of NCA is 12.2%. Overall, the proportion in CAPA level "High" who failed is 29.4%, indicating good predictive ability. Men and women in "High" have failure rates that are not significantly different from each other (women: 28.7%; men: 29.7%).

The predicted probability of NCA for each ethnic group given CAPA score is shown in Figure 10. The gray region around each line is the 95% confidence band. These confidence bands overlap across all ethnicities for all CAPA scores (except 11, since no one in the sample obtained the highest possible score).

Figure 11 shows NCA rates for released persons by ethnicity given their CAPA levels. Error bars are added to indicate the respective confidence intervals. The base rate of failure is 12.2%. Overall, the proportion in CAPA level "High" who failed is 29.4%, indicating good predictive ability. Across ethnicities, defendants in "High" have failure rates that are not significantly different from each other (visualized with the overlapping confidence intervals on the gray bars in Figure 11), indicating predictive equity.
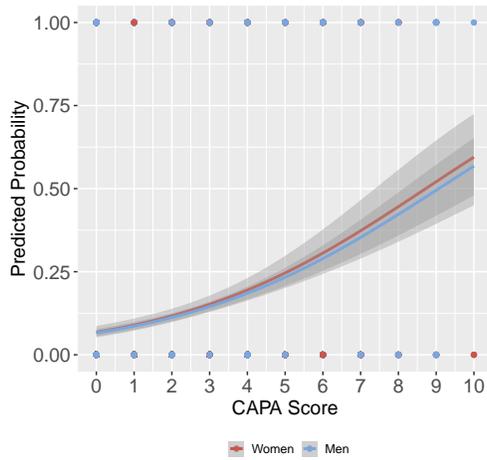
Figure 8: Predicted probabilities of NCA for CAPA score by gender.
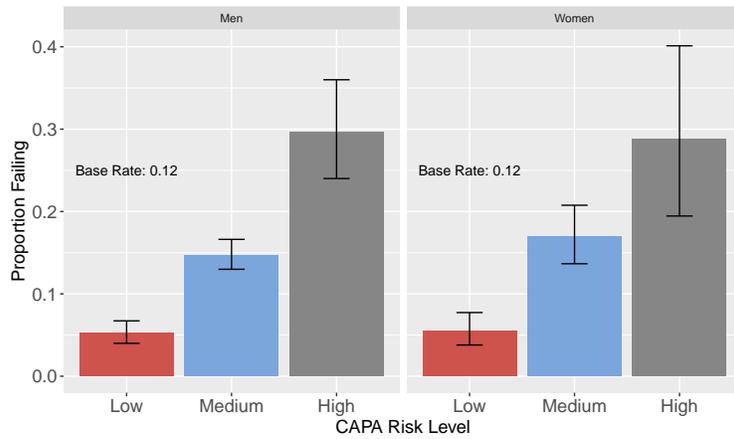
Figure 9: Proportion of defendants with an NCA on pretrial release in the levels of the CAPA by gender.
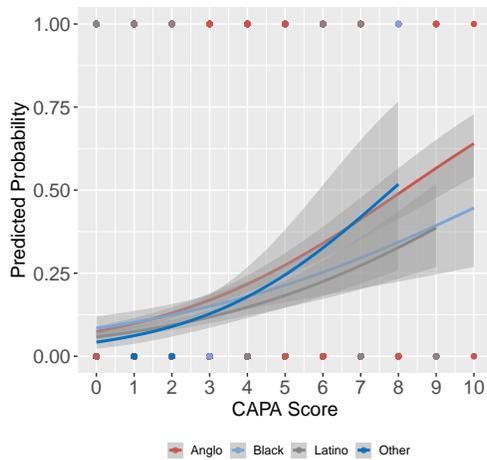
Figure 10: Predicted probabilities of NCA over CAPA score by ethnic group.
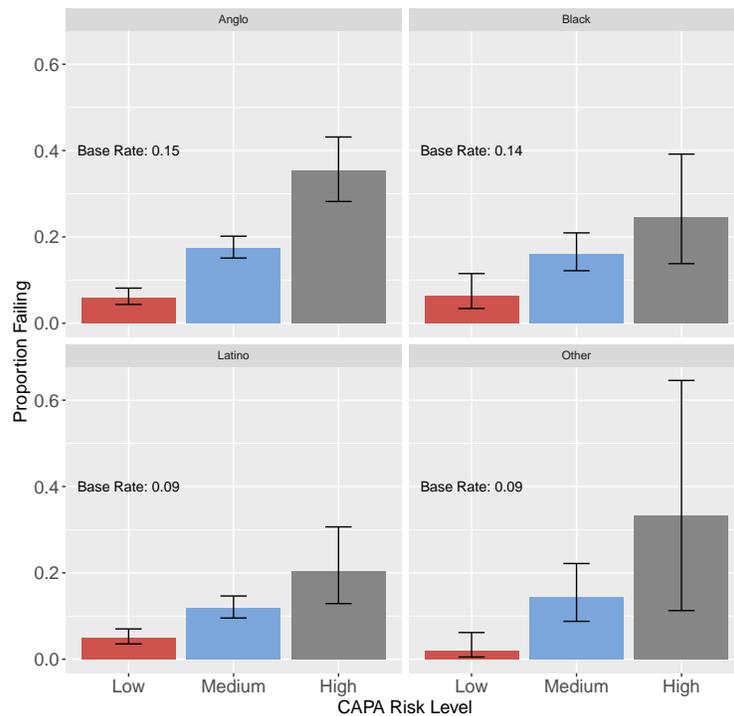
Figure 11: Proportion of defendants with an NCA on pretrial release in the levels of the CAPA by ethnic group. The 95% CI is very wide for Other in High because of the small number of failures in that level.

3.3.3 FTA Outcomes

Discriminative Ability for FTA. The AUC obtained for the CAPA in the overall sample is 0.643 (95% CI: 0.617-0.669). The AUC is 0.700 (95% CI: 0.653-0.748) for women and 0.621 (95% CI: 0.590-0.651) for men. Thus, the CAPA has fair discriminative ability in the sample overall and for men, with good discriminative ability for women. The AUCs for women and men are statistically different ($p = 0.006$), but they are not different across ethnic groups. Table 9 shows these figures for all groups, as well as the number in each subgroup, the number with FTA in each subgroup, and the percent with FTA in each subgroup.

Table 9: CAPA AUCs for FTA outcomes by overall sample and subgroups.

	AUC	Low	High	Number failing	N	Percent failing
Overall	0.643	0.617	0.669	506	3953	12.8
Women	0.700	0.653	0.748	132	1104	12.0
Men	0.621	0.590	0.651	374	2849	13.1
Anglo	0.636	0.598	0.674	234	1738	13.5
Black	0.642	0.565	0.719	58	512	11.3
Latino	0.637	0.593	0.680	187	1422	13.2
Other	0.757	0.660	0.854	27	281	9.6

Predictive Ability for FTA. The predicted probability of FTA for men and women given CAPA score is shown in Figure 12. The gray region around each line is the 95% confidence band. Since these confidence bands overlap across all CAPA scores (except 11, since no one in the sample obtained the highest possible score), we conclude that the tool predicts FTA equally well for men and women across CAPA scores.

Figure 13 shows FTA rates for released persons by gender given their CAPA levels. Error bars are added to indicate the respective confidence intervals. The base rate of failure is 12.8%. Overall, the proportion in CAPA level "High" who failed with FTA is 27.2%, indicating good predictive ability. Men and women in "High" have failure rates that are not significantly different from each other (women: 30%; men: 26.3%).

The predicted probability of FTA across ethnic groups given CAPA score is shown in Figure 14. The gray region around each line is the 95% confidence band. Across ethnicities, these confidence bands overlap across all CAPA scores (except 11, since no one in the sample obtained the highest possible score).

Figure 15 shows FTA rates for released persons by ethnicity given their CAPA levels. Error bars are added to indicate the respective confidence intervals. The base rate of FTA is 12.8%. Overall, the proportion in CAPA level "High" who failed is 27.2%, indicating good predictive ability. Across ethnicities, defendants in "High" have failure rates that are not significantly different from each other (visualized with the overlapping error bars on the gray bars in Figure 15), indicating predictive equity.

The small numbers of individuals with FTAs who are either Black (58) or Other (27) make inference for these groups unreliable (see Table 9). Typically, the number failing in any subgroup should be at least 100 in order to make valid statistical inferences. Note the wide error bars shown in Figure 15 for Blacks and Others.
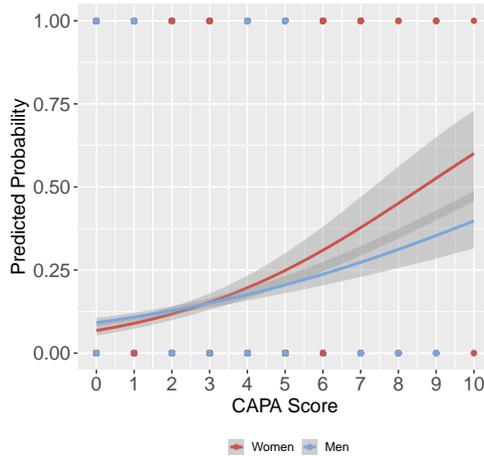
Figure 12: Predicted probabilities of FTA for CAPA score by gender.
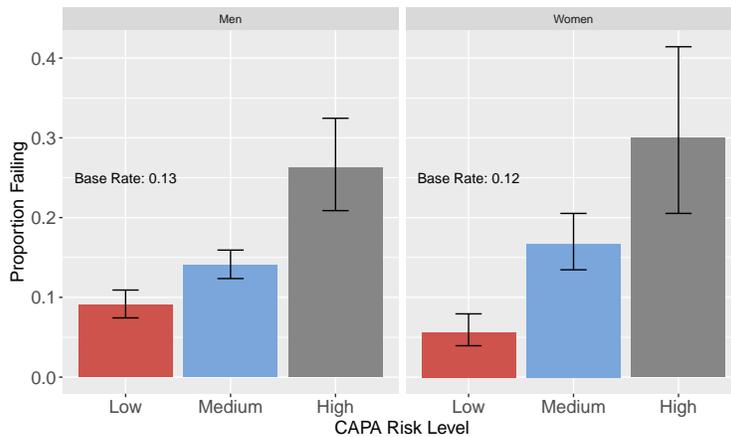
Figure 13: Proportion of defendants with an FTA on pretrial release in the levels of the CAPA by gender. The 95% CI is very wide for Other in High because of the small number of failures in that level.
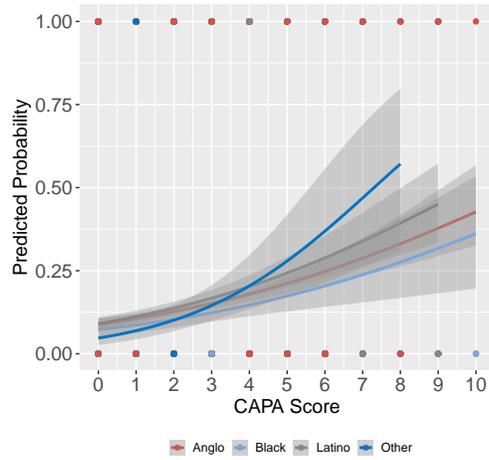
Figure 14: Predicted probabilities of FTA over CAPA score by ethnic group.
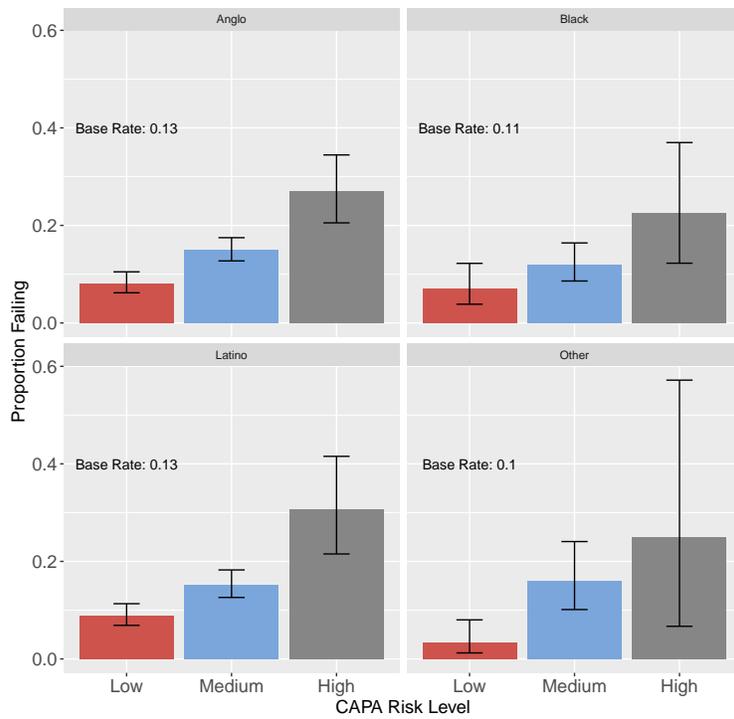
Figure 15: Proportion of defendants with an FTA on pretrial release in the levels of the CAPA by ethnic group.

3.3.4   NCA and FTA combined outcome

Discriminative Ability for aggregate NCA/FTA. The AUC obtained for the CAPA in the overall sample is 0.685 (95% CI: 0.666-0.705). The AUC is 0.726 (95% CI: 0.691-0.762) for women and 0.669 (95% CI: 0.645-0.692) for men. Thus, the CAPA has fair discriminative ability in the sample overall and for men, with good discriminative ability for women. The AUCs for women and men are statistically different ($p = 0.008$), but they are not different across ethnic groups. Thus, the CAPA has fair to good discriminative ability in the sample overall and for men and women separately, and fair discriminative ability across ethnicities. Note that those whose ethnicity was identified as Other have too few failures (47) to make reliable statistical inference about. Table 10 shows these figures for all groups, as well as the number in each subgroup, the number with either NCA or FTA in each subgroup, and the percent with either NCA or FTA in each subgroup.

Predictive Ability for NCA and FTA combined. The predicted probability of either NCA or FTA for men and women given their CAPA score is shown in Figure 16. The gray region around each line is the 95% confidence band. Since these confidence bands overlap across all CAPA scores (except 11, since no one in the sample obtained the highest possible score), we conclude that the tool predicts the aggregate outcome of either NCA or FTA equally well for men and women across CAPA scores.

Figure 17 shows the aggregate NCA/FTA rates for released persons by gender given their CAPA levels. Error bars are added to indicate the respective confidence intervals. The base rate of failure is 23.4%. Overall, the proportion in CAPA level "High" who failed with NCA/FTA is 50.3%, indicating good predictive ability. Men and women in "High" have failure rates that are not significantly different from each other (women: 52.5%; men: 49.6%).

The predicted probability of NCA/FTA across ethnic groups given CAPA score is shown in Figure 18. The gray region around each line is the 95% confidence band. Across ethnicities, these confidence bands overlap across all CAPA scores (except 11, since no one in the sample obtained the highest possible score).

Figure 19 shows NCA/FTA rates for released persons by ethnicity given their CAPA levels. Error bars are added to indicate the respective confidence intervals. The base rate of NCA/FTA is 23.4%. Overall, the proportion in CAPA level "High" who failed is 50.3%, indicating good predictive ability. Across ethnicities, defendants in "High" have failure rates that are not significantly different from each other (visualized with the overlapping error bars on the gray bars in Figure 19, indicating predictive equity).

Table 10: CAPA AUCs for FTA or NCA outcomes by overall sample and subgroups.

	AUC	Low	High	Number failing	N	Percent failing
Overall	0.685	0.666	0.705	926	3953	23.4
Women	0.726	0.691	0.762	250	1104	22.6
Men	0.669	0.645	0.692	676	2849	23.7
Anglo	0.695	0.667	0.723	454	1738	26.1
Black	0.655	0.600	0.710	121	512	23.6
Latino	0.660	0.625	0.694	304	1422	21.4
Other	0.770	0.697	0.843	47	281	16.7
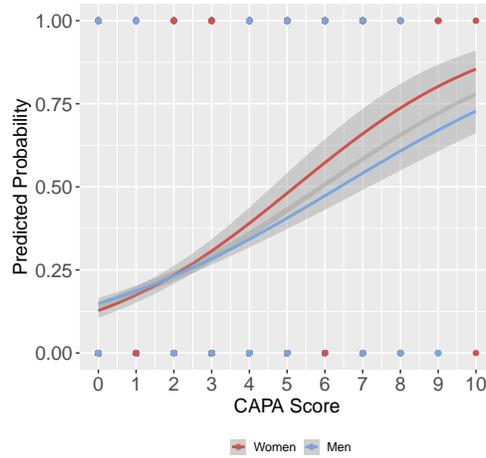
Figure 16: Predicted probabilities of FTA or NCA for CAPA score by gender.
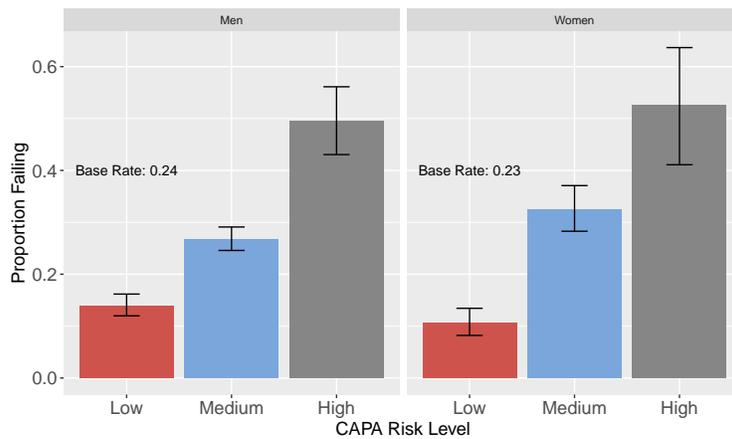
Figure 17: Proportion of defendants with an FTA or NCA on pretrial release in the levels of the CAPA by gender.
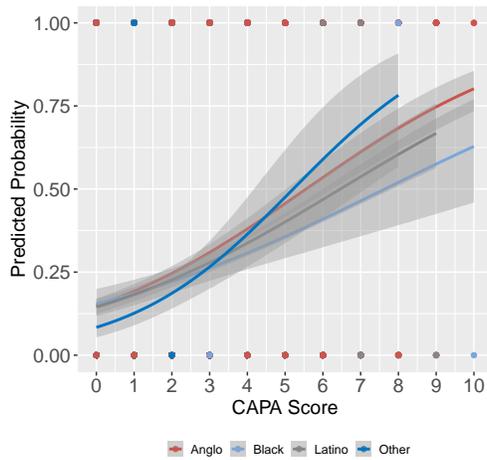
Figure 18: Predicted probabilities of FTA or NCA for CAPA score by ethnic group.
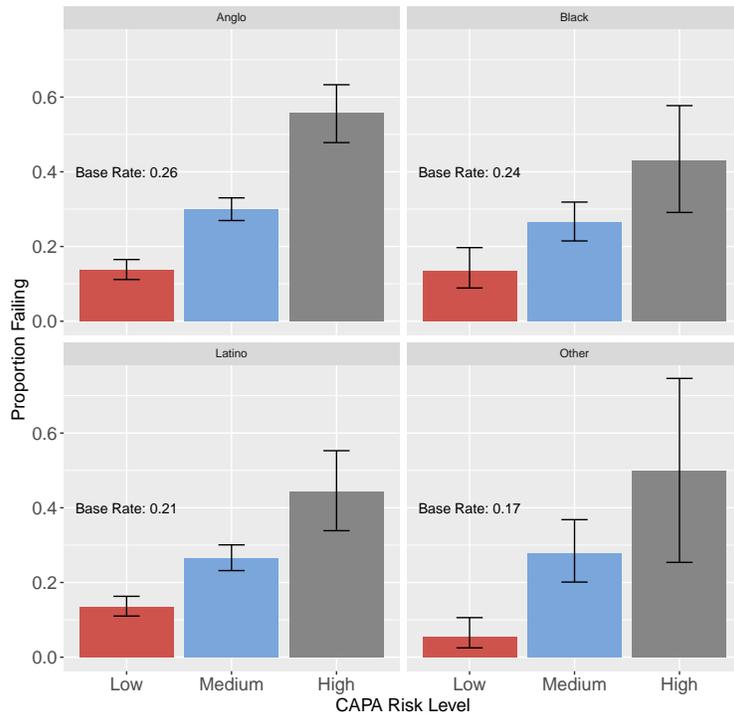
Figure 19: Proportion of defendants with an FTA or NCA on pretrial release in the levels of the CAPA by ethnic group. The 95% CI is very wide for Other in High because of the small number of failures in that level.

3.4 Cut points

The decision of where to cut the scale so that those in High fail at a substantially higher rate than the base rate, and those not in High fail at a lower rate than the base rate (or, "succeed" at a higher rate than 1 minus the base rate) can be made with the assistance of Tables 11, 12, and 13. These tables contain PPVs, Negative Predictive Values (NPVs), True Positive Fractions (TPFs), and the False Positive Fractions (FPFs), and the selection ratios, for NCA, FTA and NCA/FTA outcomes, respectively. The NPV is the proportion who do not fail given that they are not in the High risk group. The TPF is the proportion who received a High score (determined by a given cut point) given that they went on to fail, and the FPF is the proportion who received a High score (determined by a given cut point) given that they did not go on to fail. The selection ratios are the proportion in High at the given cut point.

We would like a cut point with the NPV greater than the 1 minus the base rate for that outcome, the PPV greater than the base rate, the TPF as high as possible, the FPF as low as possible, and selection ratios that are in keeping with the expectations of the agency.

Table 11 summarizes the proportions and decision metrics resulting from all possible placements of the High cut with respect to the NCA outcome. As mentioned earlier, no one in the data received the highest possible score of 11, so that possibility is not included in the table. If it were, it would be identical to what is shown for placing the High cut at 10. Placing the high cut at 6 and above allows the PPV for NCA to be 0.294, about 17 percentage points higher than the base rate of 0.122. The NPV of 0.893 is greater than 1 minus the base rate, or 0.878. There is a TPF of 0.193, meaning that of the people who eventually failed with NCA, 19.3% of them had been classified as High risk, and a FPF of 0.064, meaning that of the people who did not eventually fail with NCA, only 6.4% of them had been classified as High risk.

Table 11: Proportions and decision metrics resulting from all possible placements of the High cut with respect to the NCA outcome. The High cut was set at scores of 6 and above in this study.

	NPV	PPV	TPF	FPF	Selection Ratio (%)
>/= 0	1.000	0.122	1.000	1.000	100.00
>/= 1	0.947	0.172	0.817	0.546	57.85
>/= 2	0.936	0.211	0.682	0.353	39.33
>/= 3	0.923	0.231	0.547	0.253	28.83
>/= 4	0.914	0.254	0.441	0.179	21.09
>/= 5	0.902	0.287	0.299	0.103	12.67
>/= 6	**0.893**	**0.294**	**0.193**	**0.064**	**7.99**
>/= 7	0.885	0.297	0.091	0.030	3.74
>/= 8	0.882	0.382	0.044	0.010	1.39
>/= 9	0.881	0.462	0.025	0.004	0.66
>/= 10	0.878	0.333	0.002	0.001	0.08

Table 12 summarizes the proportions and decision metrics resulting from all possible placements of the High cut with respect to the FTA outcome. Placing the high cut at 6 and above allows the PPV for FTA to be 0.272, about 14 percentage points higher than the base rate of 0.128. The NPV of 0.885 is greater than 1 minus the base rate, or 0.872. There is a TPF of 0.17, meaning that of the people who eventually failed with FTA, 17% of them had been classified as High risk, and a FPF of 0.067, meaning that of the people who did not eventually fail with FTA, only 6.7% of them had been classified as High risk.

Table 12: Proportions and decision metrics resulting from all possible placements of the High cut with respect to the FTA outcome. The High cut was set at scores of 6 and above in this study.

	NPV	PPV	TPF	FPF	Selection Ratio (%)
$>/= 0$	1.000	0.128	1.000	1.000	100.00
$>/= 1$	0.921	0.164	0.741	0.555	57.85
$>/= 2$	0.918	0.199	0.613	0.361	39.33
$>/= 3$	0.907	0.213	0.480	0.260	28.83
$>/= 4$	0.897	0.223	0.368	0.188	21.09
$>/= 5$	0.889	0.248	0.245	0.109	12.67
$>/= 6$	0.885	0.272	0.170	0.067	7.99
$>/= 7$	0.878	0.291	0.085	0.030	3.74
$>/= 8$	0.874	0.255	0.028	0.012	1.39
$>/= 9$	0.873	0.346	0.018	0.005	0.66
$>/= 10$	0.872	0.333	0.002	0.001	0.08

Table 13 summarizes the proportions and decision metrics resulting from all possible placements of the High cut with respect to the aggregate NCA/FTA outcome. Placing the high cut at 6 and above allows the PPV for the aggregate NCA/FTA outcome to be 0.503, about 27 percentage points higher than the base rate of 0.234. The NPV of 0.789 is greater than 1 minus the base rate, or 0.766. There is a TPF of 0.172, meaning that of the people who eventually failed with either NCA or FTA, 17.2% of them had been classified as High risk, and a FPF of 0.052, meaning that of the people who did not eventually fail with either NCA or FTA, only 5.2% of them had been classified as High risk.

Table 13: Proportions and decision metrics resulting from all possible placements of the High cut with respect to the aggregate NCA/FTA outcome. The High cut was set at scores of 6 and above in this study.

	NPV	PPV	TPF	FPF	Selection Ratio (%)
$>/= 0$	1.000	0.234	1.000	1.000	100.00
$>/= 1$	0.872	0.312	0.770	0.520	57.85
$>/= 2$	0.860	0.380	0.638	0.318	39.33
$>/= 3$	0.835	0.405	0.499	0.224	28.83
$>/= 4$	0.819	0.434	0.391	0.156	21.09
$>/= 5$	0.800	0.473	0.256	0.087	12.67
$>/= 6$	0.789	0.503	0.172	0.052	7.99
$>/= 7$	0.777	0.514	0.082	0.024	3.74
$>/= 8$	0.770	0.545	0.032	0.008	1.39
$>/= 9$	0.769	0.654	0.018	0.003	0.66
$>/= 10$	0.766	0.667	0.002	0.000	0.08

Placing the High cut at 6 as we have done in this study ensures that the PPV is well above the base rate for each of the outcomes (NCA, FTA, and NCA/FTA), the NPV is greater than 1 minus the respective base rate, and that the FPF is adequately small.

CAPA scores of 1–5 are categorized as Medium risk and scores of 0 are categorized as Low risk. The score distributions observed in Section 3.2 have a substantial proportion of defendants with scores of 0. Among people with a CAPA score of 0, 94.7% did not experience an NCA event, 92.1% did not experience an NCA event, and 87.2% did not experience either an NCA or FTA event by the end of study.

4   Appendix

4.1   Additional CAPA Background

The CAPA was developed in an external sample from a 2010 pretrial study conducted in Kent County Michigan (Dieterich 2010). The development of the CAPA was influenced by California Legislation pertaining to bail reform and the use of pretrial risk assessment tools. Senate Bill 10 (SB-10) and Senate Bill 36 (SB-36) include specifications for pretrial risk assessment tools and pretrial procedure [Senate Bill 10 (2018); sb36.2019]. SB-10 was approved by the Governor and was slated to go into effect on October 1, 2019. Opponents of SB-10 were successful in adding a veto referendum to the November 3, 2020 ballot; the referendum resulted in California voters repealing SB-10. SB-36 was signed into law in November 2019.

Although SB-10 was repealed, both SB-10 and SB-36 include language that reflects ideas advanced by bail reform initiatives around the Nation. These ideas have already significantly impacted pretrial practice in California and in particular the development and use of pretrial assessments. Northpointe Research was asked to develop a replacement for the Pretrial Release Risk Scale II (PRRS-II), the Department's current pretrial assessment. The Department required an assessment that was compliant with the pretrial risk assessment tool specifications in SB-10 and SB-36.

Some of the specifications in early versions of SB-10 are not included in the version of SB-10 approved by the Governor or the final version of SB-36. Nevertheless, the ideas in early and final versions of SB-10 and the final version of SB-36 are consistent with current accepted best practices and broader trends in risk and need assessment related to fairness. These ideas stand on their own and are highly influential.

Of particular relevance to the CAPA validation study, the final version of SB-36 adds the following at Chapter 1.7. Pretrial Risk Assessment Tool Validation §1320.35.(a): "It is the intent of the Legislature in enacting this section to understand and reduce biases based on gender, income level, race, or ethnicity in pretrial release decisionmaking."

In addition, the following specifications from early and final versions of SB-10 and the final version of SB-36 influenced the development of the CAPA. Note that these specifications align very closely with the specifications of the Public Safety Assessment Court (PSA-Court)

developed by the Laura and John Arnold Foundation (LJAF). Thus the CAPA has factors in common with the PSA-Court.

- It does not require an in-person interview. (This language is not included in the final version of SB-36.)
- It does not include race, ethnicity, national origin, immigration status, gender, religion, sexual orientation, education level, employment status, socioeconomic status, arrests that did not lead to conviction, or housing status as factors used in assessing risk or determining a risk score or level. (This language was included in an early version of SB-10. The final version of SB-36 includes similar language but does not mention arrests that do not lead to conviction.)
- It is objective, standardized, and developed based on analysis of empirical data and risk factors relevant to the risk of failure to appear in court when required and risk to public safety. (Similar language is included in the final version of SB-36.)
- "Risk score" refers to a descriptive evaluation of a person's risk of failing to appear in court as required or the risk to public safety due to the commission of a new criminal offense if the person is released before adjudication of his or her current criminal offense, as a result of conducting an assessment with a validated risk assessment tool and may include a numerical value or terms such as "high," "medium," or "low" risk (version of SB-10 in the November 2020 referendum).
- Make publicly available the line items, scoring, and weighting, as well as details on how each line item is scored, for each pretrial risk assessment tool that the agency uses. (This language is in final version of SB-36.)

The following restrictions were applied during the development of the CAPA:

- Excluded *employment status* and *residence length* items from the candidate set.
- Modified the *number of arrests on bail* item to include only arrests that led to conviction.
- Excluded *age* from the candidate variable set.
- The *drug history* item is restricted to information in the case file and rap sheet obtainable without an interview.
- The regression weight scoring was replaced by summative point scoring to make the importance of the items and the scoring procedure more intuitive, transparent, and simple.
- The *top charge is felony* item was replaced in the candidate set with *top charge is felony theft or fraud* because it was anticipated that the SDSO pretrial target population would consist primarily of felony defendants if SB-10 had been enacted.

4.2   CAPA Items

Table 14 shows the items and scoring for the CAPA after updating the pilot version of CAPA. The pilot version of the CAPA was developed prior to the start of the present study in a training data set from a 2010 pretrial outcomes study conducted in Kent County Michigan. The pretrial assessment included the seven items needed to score the CAPA plus five test items that were not scored.

The pilot version of the CAPA was updated by making the following three changes:

1. dropping cc_any_arrest_on_bail_pts (Has the person been arrested/charged with a new crime that resulted in a conviction while on pretrial release?)

2. replacing cc_n_fta_pts (How many times has this person failed to appear for a scheduled criminal court hearing?) with the test item cc_n_fta_2_pts (How many times has this person failed to appear for a scheduled criminal court hearing within the last two years?)

3. adding the test item cc_n_pconviction_pts (How many times has this person been convicted of a misdemeanor or felony before as an adult?).

Table 14: Pretrial Release Assessment Modified Points Version (ScaleID = New Scale pretrialcore8)

Item	Information
cc_n_pending_pts	Number of pending charges or holds? [*PRETRIAL*] (*ToolTip:* A pending charge is a charge for which the person was previously arrested and for which a future court date is pending at the time of arrest in the current case. [**Answers:** 0=0; 1+=2]
cc_larceny_pts	Is the current top charge a felony property or fraud offense? [*PRETRIAL*] (*ToolTip:* This question refers to the top charge. The top charge is the most serious charge, as ranked by your agency, among all counts in the current pretrial case. Answer yes only if the top charge is a felony class property or fraud offense. The following are examples of property and fraud offenses: larceny, theft, theft by receiving, criminal trespass, breaking and entering, burglary, forgery, motor vehicle theft, theft at-risk person, extortion, fraud, check fraud, forgery, fraudulent use of transaction devices, arson, criminal mischief, embezzlement, money laundering, theft by deception, mail theft, counterfeiting.) [**Answers:** No=0; Yes=2]
cc_n_jail_pts	How many times has this person been sentenced to jail for 30 days or more? [*PRETRIAL*] (*ToolTip:* Include pretrial custody if 30 days or more and sentenced to time served including for violations.)
Continued on next page	

Table 14 – continued from the previous page

Item	Information
cc_n_fta_2_pts	[**Answers:** 0=0; 1=0; 2=0; 3=1; 4=1; 5+=1] How many times has this person failed to appear for a scheduled criminal court hearing within the last two years? [*PRETRIAL*] (*ToolTip:* This question refers to failure to appear for a scheduled court hearing in a criminal case within the last two years. The occurrence of prior failure to appear is typically determined by observing a prior bench warrant that was issued for failure to appear. Include the current case and all active cases.) [**Answers:** 0=0; 1=1; 2=2; 3=2; 4=2; 5+=2]
cc_drug_hx_pts	Does the person have a history of drug abuse? [*PRETRIAL*] (*ToolTip:* Consider the person's history of drug abuse of any illegal or prescription drugs, including prior drug charges and treatment for drug abuse that can be found in the person's criminal record. Do not include alcohol abuse.) [**Answers:** No=0; Yes=2]
cc_probpar_pts	Was the person on Probation or Parole at the time of the current offense? [*PRETRIAL*] (*ToolTip:* This item measures failure on parole or probation (in the community). Failure while in prison (i.e., a new offense while in prison) is a different phenomenon. In this case, the item should be answered No.) [**Answers:** No=0; Yes=1]
cc_n_pconviction_pts	How many times has this person been convicted of a misdemeanor or felony offense? [*PRETRIAL*] (*ToolTip:* Record the total number of misdemeanor or felony convictions during the person's lifetime. Do not include current case. Count one conviction per court case regardless of the number of charges associated with the case. "Convicted of an offense" means that the guilt of the defendant has been established, whether by guilty plea, trial, or plea of nolo contendere.) [**Answers:** 0=0; 1=1; 2=1; 3=1; 4=1; 5+=1]

Bibliography

Cohen, Thomas H., and Brian A. Reaves. 2004. "Pretrial Release of Felony Defendants in State Courts." Washington, D.C.: Bureau of Justice Statistics.

Desmarais, Sarah L., and Jay P. Singh. 2013. "Risk Assessment Instruments Validated and Implemented in Correctional Settings in the United States." Lexington, KY: Council of State Governments Justice Center.

Dieterich, William. 2010. "Kent County Pretrial Services Outcomes Study: Developing and Testing the COMPAS Pretrial Release Risk Scale." Traverse City, MI: Northpointe.

Guggenmoos-Holzmann, Irene, and Hans C. van Houwelingen. 2000. "The (in)Validity of Sensitivity and Specificity." *Statistics in Medicine* 19 (13): 1783–92.

Levy, Drew Griffin. 2020. "In Machine Learning Predictions for Health Care, the Confusion Matrix Is a Matrix of Confusion." Statistical Thinking; Statistical Thinking. September. https://www.fharrell.com/post/mlconfusion/.

Rhodes, William, Raymond Hyatt, and Paul Scheiman. 1996. "Predicting Pretrial Misconduct with Drug Tests of Arrestees: Evidence from Eight Settings." *Journal of Quantitative Criminology* 12: 315–48.

Royal Statistical Society Section on Statistics and the Law. 2018. "Algorithms in the Justice System: Some Statistical Issues."

San Diego Sheriff's Department. 2019. "Public Information Plan." https://sdsheriff.net/documents/public-info-2020.pdf.

Senate Bill 10. 2018. "Pretrial Release or Detention: Pretrial Services." https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB10.

Senate Bill 36. 2019. "Pretrial Release: Risk Assessment Tools." https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200SB36.

Singh, Jay P. 2013. "Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer." *Behavioral Sciences and the Law* 31: 8–22.

Steyerberg, E.W., A.J. Vickers, N.R. Cook, T. Gerds, M. Gonen, N. Obuchowski, M.J. Pencina, and M.W. Kattan. 2010. "Assessing the Performance of Prediction Models: A

Framework for Traditional and Novel Measures." *Epidemiology.* 21: 128–38.

Visher, Christy A., and Richard L. Linster. 1990. "A Survival Model of Pretrial Failure." *Journal of Quantitative Criminology* 6: 153–84.