
Kent County Pretrial Services Outcomes Study

DEVELOPING AND TESTING
THE COMPAS PRETRIAL RELEASE RISK SCALE

RESEARCH AND DEVELOPMENT DEPARTMENT

APRIL 23, 2010

Contents

Table of Contents	i
1 Executive Summary	1
2 Study Overview	4
3 Review of Pretrial Risk Instruments	6
4 Methods	13
4.1 Sample	13
4.2 Measurement	14
4.3 Analytic Approach	17
5 Results	22
5.1 Comparison of RUS and Nonsupervised Samples	22
5.2 The Association of COMPAS Items and Scales with Pretrial Failure	29
5.3 Factor Structure of Candidate Variables	33
5.4 Collinearity of Candidate Variables	34
5.5 Pretrial Risk Scale Development	36
References	50

Chapter 1

Executive Summary

This report presents results from a pretrial release outcomes study conducted in a sample of felony defendants assessed with COMPAS in Kent County, Michigan Pretrial Services. The primary objective of the study was to develop a risk scale for predicting failure to appear (FTA) and new felony arrest among defendants on pretrial release.

One purpose of pretrial release risk assessment is to sort a pretrial caseload into low-, moderate-, and high-risk groups based on the likelihood of failing to appear in court or committing a new crime pending trial. Use of the risk assessment tool by pretrial services agencies should result in consistent and equitable decisions regarding release and conditions of release. The use of objective risk assessment tools is recommended by the National Association of Pretrial Services Agencies (2004). The risk assessment tool should be empirically derived and have demonstrated predictive validity in the jurisdiction in which it is deployed. The factors that enter into the risk assessment score should be consistent with applicable state statutes.¹ These and other guiding principles for pretrial risk assessment are outlined in Pretrial Services Legal and Evidence-based Practices (VanNostrand, 2007).

Pretrial Risk Scale Development: We developed the COMPAS *Pretrial Release Risk Scale* using survival models predicting failure to appear or new felony offense pending trial in a sample of 2,831 felony defendants assessed with Core COMPAS and released pretrial. Prior pretrial risk assessment research has consistently identified a set of factors that are predictive of pretrial failure. The most common risk factors include current charges, pending charges, prior arrest history, previous pretrial failure, residential stability, employment status, community ties, and substance abuse (VanNostrand, 2003). We selected items from

¹For example in New York a pretrial risk assessment instrument cannot be based on age, gender, or marital status (Division of Probation and Correctional Alternatives, 2007).

the COMPAS assessment and included them as candidates for risk model development on the basis of this prior research.

The *Pretrial Release Risk Scale* is a weighted linear combination of risk factors (regression equation) derived through survival analysis with shrinkage applied to the weights to compensate for how the risk scale will perform when applied to a different sample. The scale has factors in common with most pretrial risk assessment tools developed for use in state courts (Cuvelier & Potts, 1993; Lowenkamp, Lemke, & Latessa, 2008; Siddiqui, 2005; VanNostrand, 2003; Winterfield, Coggeshall, & Harrell, 2003). It also has factors identified in studies of pretrial failure in federal court (VanNostrand & Keebler, 2009; Lowenkamp & Whetzel, 2009).

The *Pretrial Release Risk Scale* includes the following risk factors:

- Age at assessment
- Current top charge is felony fraud or property offense
- Number of pending charges
- Any prior arrest on bail
- Number of prior FTAs
- Numbers of times sentenced to jail
- Months residing in the community
- No stable residence
- Number of times changed residence
- Frequency of family contact
- Using drugs when arrested for current charges
- Prior alcohol abuse treatment
- Skill, trade, or profession
- Currently employed or in school

An alternate version of the risk tool was developed that excludes age at assessment, because age is not considered an acceptable objective criterion of pretrial failure in some jurisdictions.

Internal Validity: Predictive Accuracy: We evaluated the predictive accuracy of the *Pretrial Release Risk Scale* using Receiver Operating Characteristic (ROC) methods. We estimated the area under the ROC curve (AUC) in the training data and in bootstrap samples to compensate for over-optimistic results obtained in the training data. The *Pretrial Release Risk Scale* achieved an apparent AUC of .711 in the training data and an AUC adjusted for over-optimism of .688. The model with age removed had somewhat lower predictive accuracy with an apparent AUC of .694 and adjusted AUC of .673.

Pretrial misconduct is one of the most challenging outcomes to predict in a criminal justice setting. Although the predictive accuracy of the Pretrial Release Risk Scale is modest in comparison to the accuracy of many tools designed to predict more common types of outcome such as recidivism, the accuracy of the Risk Scale is on par with the best-designed pretrial risk assessment tools in the field, many of which we review in this report. The Pretrial Release Risk Scale developed in the current study has many properties that qualify it as a tool with high utility for guiding pretrial release decisions.

Limitations: We encountered challenges during the development of the Pretrial Release Risk Scale that may limit the results and how they transport to other jurisdictions. As is the case with all pretrial release samples, our study sample was subject to sample selection mechanisms. These mechanisms determined who would be assessed with COMPAS and who among that group would be released and under what conditions. These issues are discussed in the methods section of this report.

Chapter 2

Study Overview

Kent County uses an Objective Point Scale (OPS) to make bond recommendations for felony cases. The OPS yields four bond recommendation categories: (1) Personal Recognizance, (2) Personal Recognizance with Conditions, (3) Not Personal Recognizance, and (4) Preventive Detention. Approximately 5,000 felony defendants are screened with the OPS per year.

Most cases in recommendation categories 2 and 3 are also assessed with the full COMPAS at pretrial (primarily at arraignment). Between 2005 and 2008 approximately 6,763 felony defendants were assessed with the full COMPAS at pretrial. Approximately 34% of the felony defendants assessed with the full COMPAS at pretrial were subsequently released under supervision, and the remainder were released through some other mechanism (personal recognizance, bail, surety bond) or not released.

First, all felony defendants assessed with OPS between January 2005 and December 2008 were identified. Next the subset of these cases assessed with COMPAS was identified. OPS scores were provided for all pretrial felony defendants assessed with COMPAS during the period January 2005 through December 2008. The arrest date and release date (if released) that were coincident with the COMPAS assessment date were determined. These study dates were used as reference points for determining each release episode. Each released COMPAS assessed defendant contributed one episode. An episode begins on the date a defendant is first released following a felony arrest date and COMPAS assessment date.

Released Under Supervision Group: The outcomes of COMPAS-assessed defendants released under supervision were tracked in the Kent County Court Services Court View case management system. Approximately 2,300 COMPAS assessed defendants were released under supervision between January 2005 and December 2008 and had dispositions. Kent County Pretrial Services linked the COMPAS

assessment records with the pretrial case information and outcomes in Court View. The case information included the OPS items and total score, arrest date, top charge, release type (PR, PR w/Conditions, Not PR, not released, disposed at arraignment), and release date. The outcomes information included pretrial disposition and disposition date.

A defendant can experience a number of events that could qualify as pretrial misconduct. These include FTA, new arrests, and violation of supervised release conditions resulting in revocation. Technical violations include such things as failure to appear for intake appointment or comply with the conditions of supervised release. Released defendants may also violate the bond conditions that restrict contact, curfew, address change, and firearms access. The current study focuses only on failure to appear (FTA) and new felony arrest.

Other Release Group: Approximately 4,497 defendants were assessed with COMPAS between January 2005 and December 2008 but not released under supervision. The outcomes of COMPAS-assessed defendants released through other mechanisms (PR, bail, etc.) are not tracked through Court View. The outcomes for these defendants must be tracked through manual searches in a jail management system called Jail View. Local staff in Kent County that is familiar with the court conducted the manual lookup of pretrial outcomes for a random sample of these cases. An electronic pretrial outcomes tracking form was used to record the outcomes information from court case information in Jail View.

Due the cost associated with manual searches, we drew a random sample of cases from a list of all defendants assessed with COMPAS but not released under supervision. The release status of the COMPAS assessed cases in the unsupervised group were not known in advance. The defendant's case number was searched in Jail View. The release date, release type, disposition, and disposition date was recorded.

New jail bookings observed in Jail View after the release date were used to determine if there was a new arrest for a felony offense. Alternatively, if a new pretrial case was observed in Court View that was entered after the study release date and prior to the disposition in the release episode, this was used to determine if the defendant was arrested for a new felony offense. New felony offense arrests that occur outside of Kent County will not be observed in Jail View.

Chapter 3

Review of Pretrial Risk Instruments

This section provides a summary of the current state of pretrial risk assessment based on a review of reports and published studies. The following pretrial risk assessment instruments are reviewed: New York City Criminal Justice Agency Release Recommendation System [CJARRS] (Siddiqui, 2005); Virginia Pretrial Release Risk Assessment Scale [VPRAI] (VanNostrand, 2003); Urban Institute DC Pretrial Services Agency Risk Score (Winterfield et al., 2003); Harris County, Texas Bail Classification Instrument (Cuvelier & Potts, 1993), and a risk assessment instrument for U.S. Pretrial Services (Lowenkamp & Whetzel, 2009). In this section we also provide a description of the Objective Point Scale (OPS) that is used by the Pretrial Services Program in Kent County, Michigan.

New York City Criminal Justice Agency Release Recommendation System

The CJARRS was originally developed in 1974 in a sample of defendants in Brooklyn. The system has been tested and modified over the years. A current recommendation system was developed in 2003. It was validated in a pre-implementation sample of 25,278 defendants (2001) and a post-implementation sample of 20,297 defendants (November 2003 - January 2004).

The scale is constructed from the following six items:

- Does the prior bench warrant count equal zero?
- Does the open case count equal zero?

- Does defendant report a NYC area address?
- Working telephone in residence?
- Is the defendant employed or in school, or training program full time?
- Does the defendant expect someone at the arraignment?

Score ranges from -12 to +12 and results in the following four recommendation categories:

- Recommended for ROAR
- Moderate Risk for ROR
- Not Recommended for ROR
- No Recommendation (No NYSID, info only, incomplete interview)

The predictive accuracy of the CJARRS was tested using the criterion of FTA. The base rate of FTA was .16 in the pre-implementation sample and .14 in the post-implementation sample. In the pre-implementation sample, a mean cost rating (MCR) of .31 was obtained (AUC=.66). In the post-implementation sample, a MCR of .33 was obtained (AUC=.67).

Virginia Pretrial Release Risk Assessment Scale

The VPRAI was originally developed using logistic regression in a sample of 1,971 pretrial defendants that were arrested in seven localities in Virginia. The original model and a slightly revised version (warrants item dropped) were validated in two samples: A random sample of 2,778 cases from 10 pretrial service agencies that were arrested between January - December 2005 and released pretrial and a second sample that consisted of all cases released with a condition of pretrial supervision to one of 29 pretrial services agencies in Virginia. The revised scale is constructed from the following eight items:

- Two or more FTA convictions
- Current felony charge
- Any prior misdemeanor or felony conviction
- Two or more violent convictions

- Pending charges
- At current residence for at least 12 months
- Employed continuously last 24 months or primary caregiver
- History of drug abuse

The score ranges from 0 to 9 and is collapsed into 5 risk levels (low, below average, average, above average, high).

The predictive accuracy of the VPRAI was tested in the construction sample ($n=1,971$) against the criterion FTA or new offense pending disposition. The cases in the sample were tracked until final disposition. The base rate of failure was .27. In the construction sample the VPRAI obtained an AUC of .71.

In the validation conducted in the 10-agency random sample ($n=2,778$), the VPRAI obtained an AUC of .62 for FTA (base rate = .055) and an AUC of .63 for FTA or arrest (base rate = .365). In the validation in the sample of all cases released pretrial under supervision to any one of 29 pretrial services agencies in Virginia ($n=7,174$), the VPRAI achieved a AUC of .62 for FTA (base rate = .062); an AUC of .63 for FTA or arrest (base rate = .091); and an AUC of .66 for FTA, arrest, or technical violation (base rate = .32).

Urban Institute District of Columbia Pretrial Services Agency Risk Score

The Urban Institute developed a pretrial risk assessment instrument to guide pretrial release decisions in the District of Columbia. Models were trained to predict FTA and Arrest separately. The sample consisted of 7,574 first criminal cases filed between January and June 1999 that were screened by DC Pretrial Services Agency (PSA). There were 5,708 cases released pretrial. The released sample was randomly split into a construction sample and a validation sample. Regression trees (CHAID) were used to identify important interactions. Then stepwise logistic regression models were used to develop the risk models in a large set of candidate variables including the interactions identified in the regression trees. The cases were followed until the completion of PSA supervision. The Urban Institute pretrial risk instrument uses the following 22 items:

- US Citizen
- Live with family members

- Age
- How many current charges for BRA offense?
- How many current charges for obstructing justice?
- How many current charges for person offense?
- How many current charges for public order offense?
- How many current charges for property offense?
- How many charges are pending?
- How many person charges are pending?
- How many convictions (in DC)?
- How many convictions person offenses (DC)?
- How many convictions in Superior Court (DC)?
- How many charges have been filed and disposed in DC?
- How many prior times arrested in DC?
- Any arrest in jurisdiction other than DC?
- How many FTA related bench warrants in DC?
- How many invalid drug tests recorded in DC?
- How many valid tests for marijuana or hard drugs in DC?
- How many valid tests for marijuana or hard drugs submitted in last 30 days?
- How many times tested positive for hard drugs last 30 days?
- How many self-reports of marijuana or hard drug use in last 30 days?

Two subscales are calculated based on logistic regression weights: one for FTA Risk and one for Arrest Risk. The Scores ranges from 0 to 100 and are collapsed into five risk levels (Low, Condition Monitoring, Moderate, High, and Severe).

The FTA Risk Score obtained an AUC of .73 in the construction sample (base rate = .21) and an AUC of .68 in the validation sample (base rate = .21). The Arrest Risk Score obtained an AUC of .73 in the construction sample (base rate = .21) and an AUC of .71 in the validation sample (base rate = .19).

Harris County, Texas Bail Classification Instrument

The Harris County instrument was developed in a 1990 sample of 6,796 pretrial cases in the Harris County Justice Information Management Systems (JIMS). The instrument was validated in a 1993 sample of 4,710 cases that were released pretrial. Cases were followed until final disposition. Failure was defined as a new arrest or issuance of a warrant for failure to appear. The instrument was again validated in a 1995 sample of 32,589 cases released pretrial (split into construction and validation samples for recalibration and new scale development). The 1990 classification instrument was recalibrated and a new model was developed. The standard, reweighted, and alternative models were compared. The standard 1990 model was judged best and retained. The Harris County instrument is scored using the following eight items:

- Harris County Address
- Telephone
- Lives with spouse, parent or child(ren)
- Lived at address more than 1 year
- Full time employment, school, disability or homemaker
- Prior FTA
- Prior Felonies
- Prior Misdemeanors

The score ranges from - 4 to +4

The predictive accuracy of the 8-item classification instrument was tested against the criterion of FTA or arrest. In the 1990 construction sample ($n=6,796$) an AUC of .66 was obtained (base rate=.11). In the 1993 validation sample ($n=4,710$) an AUC of .65 was found (base rate = .11). In the 1995 validation sample ($n=16,294$) an AUC of .64 was found (base rate = .11).

Actuarial Risk Assessment Instrument for U.S. Pretrial Services [ARAIPS]

The ARAIPS was developed using logistic regression to predict FTA, arrest, and technical violations in a large sample of misdemeanor and felony defendants in the federal court. The sample consisted of defendants entering the federal

system between FY2001 and FY2007 who were released pretrial ($n=188,827$ to $215,338$). The sample was split into construction and validation samples. The pool of candidate variables for model development included variables identified as important risk factors for pretrial failure in work conducted by VanNostrand and Keebler (2009), as well as additional variables available in the data source. The ARAIPS is scored using the following ten items:

- Number of felony convictions
- Prior FTAs
- Pending cases
- Current offense type
- Offense class
- Age at interview
- Highest education
- Employment status
- Residence
- Current drug problems

The score ranges from 0 to 14. It is collapsed into five categories of pretrial failure risk (Category I - Category V).

The predictive accuracy of the 10-item risk instrument was tested against the criteria of FTA or arrest and FTA, arrest, or technical violation. The AUC for FTA or arrest (FTA/NCA) was .644 in the construction sample and .690 in the validation sample. The AUC for FTA, arrest, or technical violation (FTA/NCA/TV) was .726 in the construction sample and .725 in the validation sample.

Kent County Pretrial Release Objective Point Scale

The Kent County Pretrial Services uses the Objective Point Scale (OPS) to guide bail decisions and conditions of release including eligibility for pretrial services programs. The OPS is scored using items in the following eight categories:

- Type of residence

- Prior successful pretrial release
- Years residing in area
- Contact with family
- Reliability of cosigner
- Employment status
- Prior criminal history
- Probability of incarceration sentence

The OPS has a range of 0 to 45. The OPS score is used in combination with a set of classification categories (other particulars of the index offense and criminal history) to classify defendants into four categories of recommended release: (1) straight recognizance (ROR), (2) conditional release, (3) cash or surety, and (4) Preventive. Most defendants in categories 2 and 3 are routinely screened with the full Core COMPAS and are eligible for release under supervision (RUS).

Pretrial Risk Factors Identified in the State Court Processing Statistics Program

The State Court Processing Statistics (SCPS) series tracks felony cases processed in the 75 most populous counties in the U.S. The survey is conducted on a biennial basis in a sample of 40 counties selected from among the 75 most populous counties. The SCPS uses logistic regression models for clustered survey data to identify risk factors for pretrial misconduct. The survey does not include factors covering residence, employment status, community ties, mental health status, and substance abuse. However the survey does cover identified risk factors for pretrial misconduct in the following categories: release type, demographic characteristics, criminal history, and offense type. In a recent SCPS report covering results for state court felony defendants released between 1990 and 2004 several interesting findings emerged (Cohen & Reaves, 2004). Higher predicted probabilities of pretrial misconduct was associated with unsecured bond (.42) or conditional release (.37); arrest offenses for drug trafficking (.38) or larceny/theft (.37); being under age 21 (.39); being male (.35), Hispanic (.37), or black (.36); being on probation at time of the index offense (.39); having prior FTA history (.42); and having prior felony offense convictions (.39). Note the base rate of failure of any type was .33.

Chapter 4

Methods

4.1 Sample

There were 6,860 COMPAS assessments conducted with pretrial felony defendants in Kent County between January 2008 and December 2008. For cases with more than one pretrial case, the first pretrial case was retained, leaving 5,935 unique individuals with intact COMPAS assessments. Of the 5,935 unique defendants, 2,076 were released under supervision (RUS). We identified 189 tracking cases¹; 104 cases reduced prior to the first examination; 23 cases dismissed prior to the first examination, and 2 cases with missing COMPAS items. We removed these cases, leaving 1,755 unique defendants in the RUS sample.

There were 3,859 defendants assessed with COMPAS who were not released under supervision. These defendants are not tracked in Court View. Pretrial outcomes for these defendants must be searched manually in a separate system called Jail View. Due to the cost associated with manual searches, a random sample (stratified by gender) of 2,200 defendants was selected from the unsupervised group for inclusion in the study. The 189 tracking cases were added to the unsupervised sample, resulting in 2,389 defendants in the unsupervised sample. There were 736 defendants that were never released. An additional 285 defendants had their cases reduced to misdemeanors and 287 had their cases dismissed within 14 days of their arrest date (prior to the first examination). An additional 5 cases were dropped due to missing or incorrect information. This resulted in 1,076 released defendants in the unsupervised group. Thus, in the full sample there are 2,831 released defendants.²

¹A sample of defendants recommended for RUS for whom the court does not approve RUS are tracked in Court View for internal comparative analysis.

²Note that survival models were fitted in which cases dismissed or reduced prior to the

4.2 Measurement

Dependent Variables: Outcomes Measures

We define pretrial misconduct as failure to appear (FTA) or arrest for a new felony offense while on pretrial release. We develop a risk scale to predict this outcome. The decision to use this outcome as the criterion for scale development was guided by definitions of pretrial misconduct in the research literature and by common standards that we are aware of in jurisdictions around the country. For example states with statutes that do not allow preventive detention may require that the objective pretrial release decisions be made on the basis of FTA risk and not risk of public safety (new arrests). In other jurisdictions pretrial release practice standards dictate that release decisions be made on the basis of FTA and public safety. Prior research has shown that the risk factors for FTA and pretrial arrest are very similar (Goldkamp & Gottfredson, 1985; Cohen & Reaves, 2004).

If a defendant fails to appear for a scheduled appearance in the 17th Circuit Court, a bench warrant is automatically issued for the defendant. Felony offenses committed by defendants in the RUS group while on pretrial release are highly likely to be detected even if they occur outside of Kent County. Felony offenses committed by defendants in the unsupervised group may go undetected if they occur outside of Kent County.

Independent Variables: Candidate Risk Factors

An important phase of model development involved selecting variables that have been identified in prior research as risk factors for pretrial misconduct. Based on our review of prior reports and published studies, 38 variables were identified in the COMPAS data as potential candidates for model development. The candidate variables and their short descriptions are shown in Table 4.1.

Table 4.1: Candidate Variables for FTA Risk Model Development.

Items	Short Description (Response Categories)
t_prev_arrest	Total prior arrests as an adult or juvenile? (integer value)
pf_convictions	Total prior felony convictions as an adult? (integer value)

first examination were censored, as opposed to being excluded from the analysis, and the difference in the results was trivial.

Table 4.1: (continued)

Items	Short Description (Response Categories)
n.probations	How many times has person been sentenced to probation? (0=0, 1=1, 2=2, 3=3, 4=4, 5+=5)
n.arrest.on.bail	How many times arrested/charged while on pretrial release? (0=0, 1=1, 2=2, 3+=3)
n.fta	How many times failed to appear on time for court? (0=0, 1=1, 2=2, 3=3, 4=4, 5+=5)
n.jails	How many times has the offender been sentenced to jail? (0=0, 1=1, 2=2, 3=3, 4=4, 5+=5)
job	Does the person currently have a job? (1=yes, 2=no)
skill	Have a skill, trade or profession to usually find work? (1=yes, 2=no)
job_last_year	How much work or school the last 12 months? (12 Months Full Time=1, 12 Months Part Time=2, 6+ Months Full Time=3, Less Than 6 Mos. Pt/Ft=4)
wages_above_min	How hard to find a job above minimum wage compared to others? (Easier=1, Same=2, Harder=3, Much Harder=4)
havemp_r	Does person have verified local employer or school? (No=2, Yes=1)
l.drift	Do you have a regular living situation? (yes=1, no=2)
hasadd_r	Could offender provide a verifiable address? (yes=1, no=2)
yrs.address	How long have you lived at your current address? (0-5 mos.=5, 6-11 mos.=4, 1-3yrs=3, 4-5yrs=2, 6+yrs=1)
mnth.local	How long has the offender lived in this area? (0-2 mos.=4, 3-5 mos.=3, 6-11 mos.=2, 1+yrs=1)
res.moves	How often have you moved in the last twelve months? (0=0, 1=1, 2=2, 3=3, 4=4, 5+=5)
res.phone	Is there telephone at the offender's residence? (yes=1, no=2)
a_current	Using alcohol when arrested for current offense? (No=1, Yes=2)
d_current	Using drugs when arrested for current offense? (No=1, Yes=2)
ever_rx_a	Has the person ever been in treatment for alcohol?

Table 4.1: (continued)

Items	Short Description (Response Categories)
	(No=1, Yes=2)
ever_rx_d	Has person ever been in treatment for drugs?
	(No=1, Yes=2)
want_rx_a	Would benefit from treatment for alcohol?
	(No=1, Yes=2)
want_rx_d	Would benefit from treatment for drugs?
	(No=1, Yes=2)
l.fam	Do you live with family?
	(yes=1, no=2)
l.friends	Do you live with friends?
	(yes=2, no=1)
l.alone	Do you live alone?
	(yes=2, no=1)
age_at_assessment	Age at assessment
	(integer value)
age_first_arrest	Age at first arrest
	(integer value)
larceny	Is the current top charge felony property or fraud
	(yes=2, no=1)
curr.fta	Current charge- Fail to Appear
	(yes=2, no=1)
n.pending	Number of pending charges or holds
	(0=0, 1=1, 2=2, 3=3, 4+=4)
num.charges	Number of current charges
	(1=1, 2=2, 3=3, 4=4, 5=5, 6+=6)
n.rec.prob	How many times has this person had a new arrest/charge while on probation?
	(0=0, 1=1, 2=2, 3=3, 4=4, 5+=5)
has.alias	Does the offender have an alias?
	(yes=2, no=1)
fam.freq	How often last 12 mos. offender had contact with family?
	(no family=5, never=4, less than 1x/mo.=3, 1x/wk=2, daily=1)
gang_member	Are you currently a gang member?
	(yes=2, no=1)
prev_gang_mem	Have you ever been a gang member?
	(yes=2, no=1)
gang_obs	Based on screener's information is person

Table 4.1: *(continued)*

Items	Short Description (Response Categories)
	a gang member? (<i>yes=2, no=1</i>)
highgrade	What was the highest school grade attained? (<i>integer value</i>)

4.3 Analytic Approach

Predictive Modeling Strategy

Model development was guided partly by subject matter knowledge, but was primarily data-driven. We used subject matter knowledge to choose variables to include in the candidate pool. We used backward variable selection algorithms and a penalized estimation method called "least absolute shrinkage and selection operator" (LASSO) that shrinks the estimates of the regression coefficients towards zero (Tibshirani, 1996; Goeman, 2010). Some variables are shrunk all the way to zero, so the LASSO is also an alternative variable selection method.

Our analytic approach took the following steps:

- Select an initial pool of candidate variables from the COMPAS data base guided by prior pretrial studies
- Examine correlation structure and reduce the candidate pool by eliminating collinear candidate variables
- Examine nonlinear relationships and select variables using penalized (shrinkage) backward elimination
- Check the stability of the model selection procedure using bootstrap replications

Survival Models

We fitted Cox proportional hazards survival models to develop the pretrial risk scale. We fitted standard Cox models, LASSO for Cox models (Goeman, 2010), and proportional hazards competing risk models. Survival models are appropriate for these 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 & 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 & Linster, 1990; Rhodes, Hyatt, & Scheiman, 1996). In our approach we are assuming a continuous hazard of FTA over the pretrial release period, based on the empirical evidence that longer times on pretrial release increase the hazard of FTA, while recognizing that failures can only occur at scheduled court hearings. The association between time on pretrial release and the hazard for FTA is obviously a function of the number of scheduled court hearings and the types of hearing scheduled (e.g., sentencing versus other nonthreatening hearing) (Rhodes et al., 1996). But we would argue that there is also a latent tendency to FTA independent of number of court hearings that increases as the time on pretrial release increases.

The failure of interest is an FTA or arrest for a felony offense while on pretrial release. Survival time begins on the release from pretrial detention date. Survival time is measured in days from release date to the point of the first failure of interest. Cases can experience the event of interest as well as competing events. Pretrial release may be revoked for a technical violation, the defendant's case could be dismissed, the defendant's case could be reduced to a misdemeanor and transferred to the lower court, or the defendant could successfully complete the pretrial release period without any misconduct. These competing events alter the probability of observing the event of interest. For example, if the defendant's pretrial release is revoked for a technical violation and they are returned to jail, they cannot fail by the event of interest (FTA or felony arrest). We fit two types of survival model: cause-specific and competing risk regression. In the cause-specific survival models cases that experience a competing event are censored and removed from the risk set. In the competing risk survival models the hazard of the event of interest (FTA or felony arrest) is adjusted by the survival probability of the competing events. In the competing risk models, cases that are reduced and transferred are censored at that point, as opposed to being treated as a competing event.

- Event of interest: FTA or felony arrest.
- Competing events: death³, case dismissed, release revoked for technical violation, or successful completion of pretrial release.
- Censoring: Case reduced and transferred to district court or case still pending on study end date (November 30, 2010).

³Four RUS defendants died while on pretrial release. These events were added to the dismissed category.

Variable Transformations

We used a variable selection approach that simultaneously tests nonlinear relationships between each candidate variable (continuous variables only) and the study outcome (Sauerbrei & Royston, 1999). We used multiple fractional polynomial (MFP) functions in STATA and R. Variables are selected using backward elimination (Cox regression) to identify the best subset of the candidate variables to predict each outcome. The nominal p -value was set at .10. In backward elimination all variables are first entered into the model. In a stepwise manner each variable in turn is removed from the model and the effect of its removal is tested. A variable is dropped from the model if its removal causes a nonsignificant (p -value > .10) increase in the deviance. After identifying the best fitting polynomial transformation all variables are centered at their respective means to clarify interpretation of the coefficients. Stable variable transformations identified by MFP were applied to some of the candidate variables for the Cox LASSO models as well. The MFP approach works best with continuous variables. Most of our candidate variables are discrete, thus MFP had a limited role in our model development.

Model Stability

We tested the stability of the variable selection using bootstrap methods. The bootstrap method refits the model a large number of times ($n=200$) in resamples of the study data and keeps track of the fraction of times each variable is selected in a bootstrap resample. This is one method that we used to test the internal validity of model development. We focus on internal validity tests because we did not have enough data for a holdout sample that would have allowed us to test the external validity of the model we developed. We also evaluated the over-optimism of our estimates of predictive accuracy (see next).

Measures of Association

We use survival models to estimate three measures that are useful for evaluating the predictive value of the risk scales: failure probabilities, hazard ratios, and the area under the receiver operating characteristic curve (AUC).

- *Failure Probability.* In typical survival data without competing events the Kaplan-Meier statistic is used as an estimate of survival or failure probability (1 - KM) at different time points. However, in the context of competing risks, 1 - KM is not a proper or interpretable failure

probability. Competing events also violate the independence assumption that was discussed above in the context of informative censoring (Putter, Fiocco, & Geskus, 2007). In this report we calculate the crude cumulative incidence function. This gives the probability of failure adjusted for failures from competing events. We also use a specialized proportional hazards survival model for competing risks (Fine & Gray, 1999) to estimate and plot failure probabilities (cumulative incidence curves) for each type of failure within the levels of the risk scores. The goal is to compare the **probability of the event of interest** (e.g., arrest) between the levels of the risk scores. The generality of the results from competing risk models are limited to populations with similar characteristics and similar patterns of competing events (Pintilie, 2007).

- *Hazard Ratio.* We use the Cox survival model to model the effect of the risk scales on the cause-specific hazard for any arrest, person offense arrest, absconding, and new commitment. In these models failures due to competing events are censored. The cause-specific hazards reflect the risk of the event of interest (e.g., arrest) as if the competing events did not exist. The goal is to compare the **hazard rates for the event of interest** within different levels of the risk scores. The results are valid for any population with similar characteristics regardless of the pattern of competing events (Pintilie, 2007).
- *Area Under the Curve.* To gauge the predictive accuracy of the pretrial misconduct risk scale, we estimated the area under the receiver operating characteristic curve (ROC). We refer to this measure as the area under the curve (AUC). For survival models, the concordance index is equivalent to the AUC. The concordance index is defined as the probability that the predictor values and survival times for a pair of randomly selected cases are concordant. A pair is concordant if the case with the higher predictor value has a shorter survival time. The calculation is based on the number of all possible pairs of non-missing observations for which survival time can be ordered and the proportion of relevant pairs for which the predictor and survival time are concordant (F. E. Harrell, Califf, Pryor, Lee, & Rosati, 1982). The AUC is a good method for evaluating prognostic models because the estimate is not influenced by the base rate (proportion of the sample that fails). This characteristic makes it easier to compare results across studies. However, the base rate can affect the precision of the AUC. The AUC ranges from .50 to 1.00. An AUC of .50 indicates no relationship between the risk scale and the outcome. An AUC of 1.00 indicates that the risk scale predicts the outcome perfectly. In the pretrial risk assessment literature that we reviewed the reported AUCs ranged from .64 thru .74.

We assessed the amount of over-optimism in our AUC estimates by using a bootstrap routine described by [F. E. Harrell Jr. \(2001\)](#). The approach proceeds as follows:

- Draw a bootstrap sample of size n with replacement from original sample of size n
- Derive a model in the bootstrap sample and calculate prediction equation XB and get AUC
- Apply model XB to the original sample and calculate AUC
- Subtract AUC in bootstrap sample from AUC obtained by applying XB to original sample to estimate optimism
- Repeat for 200 or so bootstrap replications, obtain average over-optimism and subtract from apparent AUC

© 2010 Northpointe, Inc., d.b.a. equivalent. All rights reserved.

Chapter 5

Results

5.1 Comparison of RUS and Nonsupervised Samples

In this section we discuss sample selection bias and compare the RUS and unsupervised samples. The outcomes for the RUS sample are collected in a standardized manner in a case management system called Court View developed specifically for this purpose. On the other hand, the outcomes of the unsupervised release sample were collected by manually searching records in a jail management system called Jail View. Thus the two groups may be distinct in more ways than one. It is for this reason that we compare the two samples in terms of the candidate variables and outcomes. We also develop models separately in the RUS sample and in the combined sample. Note that we did not have a sufficient sample of released unsupervised defendants to develop a separate model in that sample.

Sample Selection Bias

As is the case in all pretrial release samples, the Kent County pretrial study sample is affected by selection mechanisms that determine which defendants are released and included in the estimation sample. More than one selection mechanism is operating on the study sample. To be included in the study sample, defendants must have an OPS score of 3 or 4. Then they must be assessed with the COMPAS. Next they had to be released, either into the supervised release program or through some other mechanism (bail/bond, PR). The study sample excludes cases with OPS scores of 1 (lower risk cases recommended for release on personal recognizance) and cases with OPS scores of 4 (higher risk cases recommended for preventive detention). The results

that we obtain in the study sample of felony defendants may not generalize to other settings or other types of pretrial defendant.

Pretrial Release Types

In the total sample of COMPAS assessed defendants ($n=4,265$), valid OPS scores were available for 4,163 defendants. Among those defendants with a valid OPS score, 1,436 were assigned to category 2 (Personal Recognizance with Conditions) and 2,727 were assigned to category 3 (Not Personal Recognizance).

Of those defendants in category 2, there were 776 released under supervision (RUS) and 660 released through some other mechanism or not released. Among the 776 RUS defendants, 348 were released on bail/bond and 428 were released on personal recognizance. Of the 660 defendants in the unsupervised group, 289 bonded out, 51 were released on personal recognizance with conditions, 85 were released on personal recognizance, 92 were never released, and 143 had their cases dismissed or reduced to misdemeanors within 14 days of their arrest date (prior to the preliminary examination).

Of the defendants in category 3, there were 1,065 released under supervision (RUS) and 1,662 released through some other mechanism or not released. Among the 1,065 RUS defendants, 708 were released on bail/bond and 357 were released on personal recognizance. Of the 1,662 defendants in the unsupervised group, 470 bonded out, 100 were released on personal recognizance with conditions, 51 were released on personal recognizance, 635 were never released, and 406 had their cases dismissed or reduced to misdemeanors within 14 days of their arrest date (prior to the preliminary exam).

The RUS group contains a smaller fraction of cases with category 3 OPS scores (58%) compared to the unsupervised group (72%). However, among cases actually released pretrial, similar percentages of defendants in the RUS group (58%) and the unsupervised group (60%) have category 3 OPS scores.

Among released defendants a smaller percent of RUS defendants posted bail/bond (57%) compared to unsupervised defendants (72%). Only 14% of released defendants in the unsupervised sample were released with conditions, while all RUS defendants are released with some type of condition.

OPS Overrides

There are 3,995 COMPAS assessed defendants with both a valid OPS score and a valid Pretrial Services bond recommendation. Table 5.1 compares the OPS score with the Pretrial Services recommendation. Pretrial Services overrode 19.9% of the OPS scores up from category 2 to category 3 and 8.8% of the OPS scores down from category 3 to category 2. The Pretrial Services recommendation is made to the court. The final pretrial release decision is made by the court.

Table 5.1: Comparison of OPS Score and Pretrial Services Bond Recommendation ^a

	PT 2	PT 3	Total
OPS 2	1,076 80.1%	268 19.9%	1,344 100.0%
OPS 3	233 8.8%	2,418 91.2%	2,651 100.0%
Total	1,309 32.7%	2,686 67.3%	3,995 100.0%

^a Cases with a valid OPS score and pretrial services recommendation.

Top Charges

The top charges for the RUS sample and the unsupervised sample are shown in Table 5.2 Table 5.3, respectively. A larger proportion of the RUS sample has a top charge of drug trafficking (.21) compared with the unsupervised sample (.16). For a somewhat larger proportion of the unsupervised sample, the top charge is a property offense (.17), compared with the RUS sample (.14). Otherwise the two samples are very similar in terms of top charge. It must be kept in mind that these comparisons are for cases that were released. Among all defendants assessed with COMPAS there are a few notable differences in top charge between the RUS and unsupervised samples (results not shown). In the broader sample of all defendants assessed with COMPAS, a larger proportion of the nonsupervised defendants have a top charge of assault (.20) compared to the RUS defendants (.15). In the broader sample there are also slightly higher proportions of unsupervised defendants with coercive sex offenses and robbery, an effect which is opposite of that observed among the released defendants. In

general however we don't see many differences in terms of top charge between the two samples.

Table 5.2: Top charge for defendants in the supervised release sample.

offense group	frequency	relative frequency
Arson	1	0.00
Assault	237	0.14
Burglary	84	0.05
DrugP	236	0.14
DrugT	369	0.21
DUI	100	0.06
Fraud	140	0.08
Missing	9	0.00
Other	149	0.08
Property	251	0.14
Robbery	33	0.02
SexForce	29	0.02
SexWOForce	9	0.00
Weapon	108	0.06
Sum	1,755	1.00

Dispositions

The dispositions were collapsed into six categories called events. The events for the RUS sample are shown in Table 5.4. Success is defined as making it through the pretrial release period without any misconduct. The category labeled "Reduced" includes defendants who had their cases reduced to misdemeanors and transferred to a lower District Court. The reduced and dismissed events occurred after the first examination. Table 5.5 shows the dispositions for defendants in the unsupervised release sample collapsed into the six categories.

Competing Events

Figure 5.1 shows a plot of the cumulative crude incidence function of FTA or felony arrest for released defendants in the full sample. The cumulative crude incidence is the failure probability in the presence of the competing events case dismissal, technical violation, and successful completion of pretrial release.

Table 5.3: Top charge for defendants in the non-supervised release sample.

offense group	frequency	relative frequency
Assault	177	0.16
Burglary	39	0.04
DrugP	130	0.12
DrugT	176	0.16
DUI	63	0.06
Fraud	104	0.10
Other	127	0.12
Property	187	0.17
Robbery	13	0.01
SexForce	13	0.01
SexWOForce	4	0.00
Weapon	43	0.04
Sum	1,076	1.00

Table 5.4: Events for defendants in the RUS sample.

Disposition	frequency	relative frequency
Reduced	117	0.07
FTA	177	0.10
Felony Arrest	40	0.02
Technical	129	0.07
Dismissed	93	0.05
Success	1,199	0.68
Sum	1,755	1.00

Table 5.5: Events for defendants in the unsupervised release sample.

Disposition	frequency	relative frequency
Reduced	262	0.24
FTA	50	0.05
Felony Arrest	19	0.02
Technical	19	0.02
Dismissed	47	0.04
Success	679	0.63
Sum	1,076	1.00

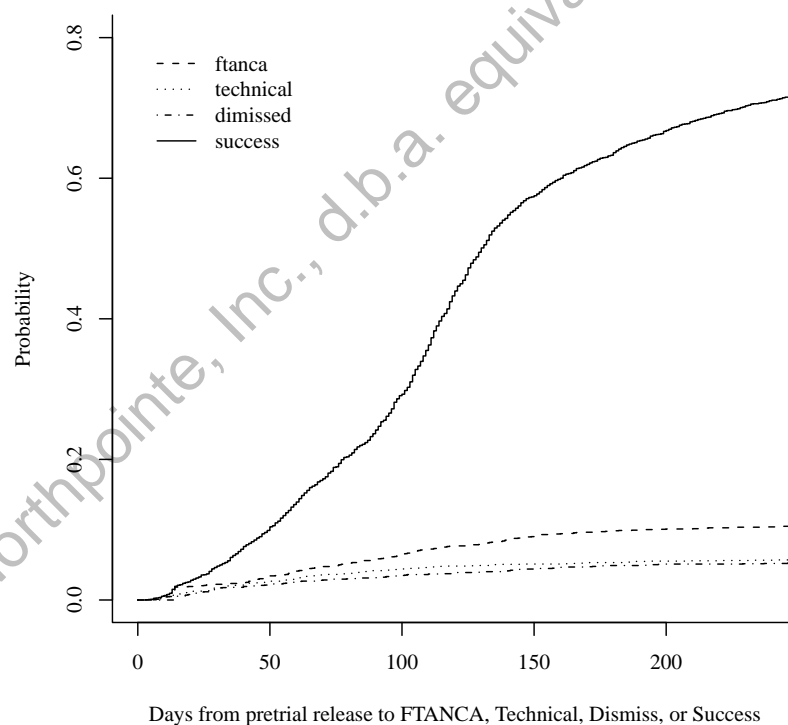


Figure 5.1: Plot of the probability (cumulative incidence) of FTA or felony arrest in the presence of competing events technical violation, dismissed, and success ($n=2,831$).

Comparison of the COMPAS Candidate Variables in the RUS and Unsupervised Samples

Table 5.6 shows the results of tests of the difference of COMPAS item means in the RUS and unsupervised release samples. The unsupervised release sample has a more serious profile as evidenced by higher mean scores on prior arrests (`t_prev_arrest`), felony convictions (`pf_conviction`), probations (`n_probations`), new charges on probation (`n_rec_prob`), arrests on bail, and FTAs. A higher percent of the RUS sample was using drugs when arrested on the current charges (`d_current`). Note that for the mean of dichotomous items such as `d_current` (`no=1`, `yes=2`), the percent scoring 2 is obtained by subtracting 1 from the mean. For example, 21% of the unsupervised sample and 26% of the RUS sample were using drugs when arrested on the current charges.

Table 5.6: Results from t-tests comparing the means for each candidate COMPAS question in the supervised release and other release samples.

item	supervised	unsupervised	t-value	p-value
male	0.84	0.81	2.04	0.041
agetnrc	32.26	31.92	0.78	0.433
age_first_arrest	20.44	20.70	-1.02	0.308
t_prev_arrest	9.56	8.61	2.75	0.006
pf_conviction	1.57	1.58	-0.06	0.954
n_probations	1.41	1.48	-1.60	0.111
curr_fta	1.04	1.04	-0.20	0.843
larceny	0.22	0.27	-3.17	0.002
n_pending	0.11	0.12	-0.76	0.449
num_charges	1.59	1.59	0.08	0.937
n_rec_prob	0.22	0.27	-2.49	0.013
has_alias	1.14	1.12	1.69	0.090
n_arrest_on_bail	0.19	0.21	-0.99	0.324
n_fta	1.08	0.88	3.95	0.000
n_jails	2.60	2.44	2.77	0.006
fam_freq	1.30	1.32	-0.59	0.553
l_alone	1.13	1.13	0.15	0.879
l_fam	1.32	1.31	0.84	0.401
l_friends	1.18	1.16	1.32	0.186
l_drift	1.06	1.07	-0.77	0.441
hasadd_r	1.05	1.06	-0.91	0.362
yrs_address	3.29	3.26	0.55	0.583
mnth_local	1.28	1.32	-1.54	0.124

Table 5.6: *(continued)*

item	supervised	unsupervised	<i>t</i> -value	<i>p</i> -value
res_moves	0.70	0.74	-0.87	0.383
res_phone	1.15	1.14	0.50	0.617
gang_member	1.01	1.01	0.81	0.418
prev_gang_mem	1.04	1.03	1.12	0.263
gang_obs	1.00	1.01	-1.12	0.264
a_current	1.30	1.27	1.62	0.106
d_current	1.26	1.21	3.48	0.001
ever_rx_a	1.28	1.30	-0.92	0.359
ever_rx_d	1.32	1.31	0.65	0.513
want_rx_a	1.25	1.23	1.18	0.239
want_rx_d	1.31	1.26	2.85	0.004
job	1.58	1.57	0.64	0.522
skill	1.70	1.67	1.77	0.076
job_last_year	3.36	3.27	1.79	0.074
wages_above_min	2.35	2.32	0.79	0.431
havemp_r	1.58	1.56	0.67	0.504
highgrade	11.44	11.43	0.18	0.855

5.2 The Association of COMPAS Items and Scales with Pretrial Failure

In this section we report the results from univariable Cox survival models for combinations of outcomes and sample. Table 5.7 shows the results of fitting separate survival models for each of candidate variables from Core COMPAS. The table shows the *t*-values from the models. If the *t*-value is less than 2.00, this indicates that the relationship between the item and the pretrial outcome is not significantly different than zero. Values less than 2.00 are represented by blanks in the table to make it easy to see which scales have significant bivariate associations with the pretrial outcomes.

The results indicate that most of the candidate variables have significant bivariate associations with all the pretrial outcomes in the combined and RUS samples. A few of the items do not demonstrate associations with any of the outcomes. A few examples include prior felony convictions (pf_convictions), current fta (curr_fta), number arrests on probation (n_rec_prob), has alias (has_alias), the living situation items (l_alone, l_fam, l_friends), the gang affiliation items (gang_member, prev_gang_mem, gang_obs), and using alcohol when arrested (a_current). Interestingly only a few items have associations

with the outcomes in the unsupervised sample. In general the associations are stronger in the combined sample due primarily to an increase in sample size, on which the t -values depend. Similarly the nonsignificant results in the unsupervised sample are a function of the small number of FTAs (50) and felony arrests (19) in the unsupervised sample. For survival models the effective sample size is equal to the number of failures.

Out of a concern to develop a parsimonious model that could be completed with a minimum of items, we decided not to include existing scales as candidates for developing the FTA Risk Scale. However, for the sake of completeness and because we have an interest in learning about the relationship between the domains covered by the COMPAS Scales and the pretrial outcomes, we fitted separate survival models for each Core COMPAS scale and present the results in Table 5.8. Again the table shows the t -values from the models with values less than 2.00 suppressed. The results indicate that all the COMPAS scales have significant bivariate associations with all the pretrial outcomes in the combined and RUS samples, except History of Violence (HistViol), Financial Problems (Financ), and Family Crime (FamCrim). Pretrial release agencies in some jurisdictions use the results from the Core COMPAS assessment to guide the selection of conditions and interventions for defendants released under supervision or with other conditions. These results support the use of the Core COMPAS assessment for these purposes.

Table 5.7: Results (t -values) from univariable cause-specific Cox models regressing the hazard of any pretrial misconduct and FTA or felony arrest on the COMPAS items in the combined, unsupervised, and supervised samples.

Item	Any Misconduct			FTA or Felony Arrest		
	Comb	Unsuper	Super	Comb	Unsuper	Super
male						
age <trnc< td=""><td>−2.60</td><td></td><td>−3.12</td><td></td><td></td><td>−2.39</td></trnc<>	−2.60		−3.12			−2.39
age_first_arrest	−2.18		−2.32			
t_prev_arrest	3.79		3.23	3.22		2.44
pf_convictions						
n_probations						
curr_fta					2.38	
larceny	6.37	4.07	6.03	6.36	4.79	5.20
n_pending	3.01		2.89	3.61		3.15
n_rec_prob						
has_alias						
n_arrest_on_bail	3.57	3.66	2.35	3.11	4.92	

Table 5.7: *(continued)*

	Comb	Unsuper	Super	Comb	Unsuper	Super
n_fta	4.36		3.40	3.75	2.23	2.77
n_jails	3.61		2.94	3.22		2.54
fam_freq	5.41		5.32	5.31		5.27
Lalone						
Lfam						
Lfriends						
Ldrift	3.90		3.49	3.83		3.11
hasadd_r	3.43		2.39	3.43	2.54	2.02
yrs_address	2.99		2.41	3.73		3.03
mnth_local	3.38		3.79	4.50		4.86
res_moves	4.42	2.99	3.29	4.90	3.18	3.64
res_phone	2.83			2.25		
gang_member						
prev_gang_mem						2.22
gang_obs						
a_current						
d_current	4.37		4.09	3.34		3.01
ever_rx_a			2.43			2.40
ever_rx_d	2.69		2.15	2.54	2.16	
want_rx_a			2.18			
want_rx_d	4.20		3.88	2.99		2.43
job	2.65		3.58	3.16		3.85
skill	3.36		4.31	3.04		4.01
job_last_year	3.14		3.12	3.07		3.08
wages_above_min			2.71	2.17		3.12
havemp_r	2.81		3.67	3.14		3.74
highgrade	-2.94		-3.43			

Table 5.8: Results (t -values) from univariable cause-specific Cox models regressing the hazard of any pretrial misconduct and FTA or felony arrest on the COMPAS scales in the combined, unsupervised, and supervised samples.

Item	Any Misconduct			FTA or Felony Arrest		
	Comb	Unsuper	Super	Comb	Unsuper	Super
CrimInv	2.81			2.51		2.33
HistNonC	3.92		2.39	3.31	2.51	3.44
HistViol						
CurrViol	-3.15	-2.54	-2.61	-2.88	-2.11	-2.80
CassPeer	4.08	2.24	2.66	3.50	2.40	3.36
SubAbuse	2.83		2.44	2.37		3.32
Financ						2.49
VocEd	4.30		5.00	4.13		5.27
FamCrim						
SocEnv	3.27		3.14	3.90		2.46
Leisure	3.08		3.49	3.21		3.60
ResInst	5.00	2.15	4.68	5.60	2.62	4.26
SocAdj	4.42		4.12	3.65		4.91
EJuvSoc	3.49		3.59	2.99		4.09
CrimOpp	6.67		6.83	6.75		6.85
Soc.Isolation	3.26		2.77	3.34		2.92
CrimAttC	2.79		2.58	2.98		2.28
CrimPers	4.92		3.73	3.95		4.77

Note. t -values less than 2.00 are not printed.

5.3 Factor Structure of Candidate Variables

We examined the factor structure of the candidate variables to evaluate their suitability to the task of developing a pretrial misconduct risk scale. Table 5.9 shows the results of a factor analysis of the initial pool of candidate variables from the COMPAS data base. The pattern matrix is from a factor solution with oblimin rotation with allows the derived factors to correlate. The rotation makes the factor structure more apparent. Factor loadings less than .20 are printed as blanks in the table. A six-factor solution was obtained. Four well-defined factors emerged covering the following domains: employment, criminal involvement, residential stability, and substance abuse. The factor solution indicates that the candidate variables represent the most important pretrial misconduct risk factors identified in prior research (VanNostrand & Keebler, 2009).

Table 5.9: Pattern matrix from a factor analysis of the candidate variables for the FTA/NCA risk model in the combined sample.

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
t_prev_arrest	0.944					
pf_conviction	0.591					
n_probations	0.415					
n_arrest_on_bail	0.465					
n_fta	0.675					
n_jails	0.787					
job		0.987				
skill		0.491				
job_last_year		0.604				
wages_above_min		0.304				
havemp_r		0.990				
Ldrift			0.789			
hasaddr			0.759			
yrs_address			0.510			
mmth_local			0.416			
res_moves			0.717			
res_phone			0.521			
a_current				0.514		
d_current				0.395		
ever_rx_a				0.563		
ever_rx_d				0.446		
want_rx_a				0.736		

Table 5.9: (continued)

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
want_rx_d				0.559		
lfam					0.669	0.593
lfriends					1.007	
lalone						0.926
agetrc	0.266			0.220		0.242
age_first_arrest						
larceny						
curr_fta						
n_pending						
num_charges						
n_rec_prob						
has_alias	0.271					
fam_freq			0.286		0.219	
gang_member						
prev_gang_mem						
gang_obs						
highgrade						

5.4 Collinearity of Candidate Variables

The results from variable selection methods such as backward elimination can be influenced by highly correlated variables. This is because highly correlated variables can compete with each other for inclusion in the model. The correlation matrix of the candidate variables can be examined to identify pairs of variables with high correlations (for example, Pearson r greater than .60). Candidate variables involved in collinear relations will also show up in principal components analysis as short principal components with small eigen values and large loadings. Collinearity among the candidate variables also can be measured by the variance inflation factor (VIF). The VIF is based on the multiple correlation from the regression of each candidate variable on all the other candidate variables (predict each variable with all other candidate variables). A VIF greater than 4 indicates the variable is involved in a collinear relationship. The results of the stability analysis (described in Section 5.5) can also be examined to identify pairs of variables that compete for inclusion in the model (for example, variable x_1 is included in the model only when x_2 is excluded). We employed all these methods to address the problem of collinearity. The result of the VIF analysis is shown in 5.10. As shown, based on the VIF we can identify five variables

that are involved in collinear relations : job:havemp_r and Lfam:Lfriends:L-alone. We made a decision to drop 'job' and 'Lalone' from the model selection phase of risk scale development. Doing so reduced the VIF for all remaining candidate variables.

We dropped want_rx_a, want_rx_d, gang_member, and prev_gang_mem for substantive reasons (item reliability and interpretation) and because these items were from domains covered by other items in the candidate pool that were more interpretable and reliably measured.

Table 5.10: Variance inflation factors (VIF) for the initial candidate variables in the combined sample.

Candidate	VIF
t_prev_arrest	3.71
pf_conviction	1.85
n_probations	1.67
n_arrest_on_bail	1.74
n_fta	1.93
n_jails	2.73
job	24.11
skill	1.59
job_last_year	1.81
wages_above_min	1.30
havemp_r	24.46
Ldrift	2.04
hasadd_r	1.92
yrs_address	1.96
mnth_local	1.26
res_moves	2.31
res_phone	1.34
a_current	1.49
d_current	1.62
ever_rx_a	1.97
ever_rx_d	1.90
want_rx_a	2.09
want_rx_d	2.20
Lfam	15.57
Lfriends	10.26
Lalone	8.25
age_first_arrest	1.17
larceny	1.06
curr_fta	1.09

Table 5.10: *(continued)*

Candidate	VIF
n_pending	1.33
num_charges	1.10
n_rec_prob	1.25
has_alias	1.10
fam_freq	1.30
gang_member	1.37
prev_gang_mem	1.32
gang_obs	1.08
highgrade	1.15

5.5 Pretrial Risk Scale Development

Variable Transformations and Selection Methods

We used a variable selection approach that simultaneously tests nonlinear relationships between each candidate variable (continuous variables only) and the study outcome (Sauerbrei & Royston, 1999). We used multiple fractional polynomial (MFP) functions in STATA and R. Variables are selected using backward elimination (Cox regression) to identify the best subset of the candidate variables to predict each outcome. The nominal p -value was set at .10. In backward elimination all variables are first entered into the model. In a stepwise manner each variable in turn is removed from the model and the effect of its removal is tested. A variable is dropped from the model if its removal causes a nonsignificant (p -value > .10) increase in deviance. After identifying the best fitting polynomial transformation all variables are centered at their respective means to clarify interpretation of the coefficients.

We tested the stability of the variable selection using bootstrap methods. The bootstrap method refits the model a large number of times ($n=200$) in resamples of the study data and keeps track of the fraction of times each variable is selected in a bootstrap resample.

Note that the polynomial transformations are applied only to continuous level candidate variables. For the purpose of this study we define a continuous variable as a numeric variable having a range of 6 or greater, which is the default behavior in the MFP package. In all models age-at-assessment was the only variable for which a polynomial transform improved the fit over a linear term.

We also used a penalized estimation method called the "least absolute shrinkage and selection operator" (LASSO) that shrinks the estimates of the regression coefficients towards zero (Tibshirani, 1996; Goeman, 2010). Some variables are shrunk all the way to zero, so the LASSO is an alternative variable selection method.

Fitting the Models and Deriving the Pretrial Release Risk Scale

We developed two model versions, one that included age and one that excluded age. This was done to provide a version of the pretrial risk score that could be used in jurisdictions that exclude age, gender, marital status, or other demographic factors from consideration when making pretrial release decisions.

For each model version, separate risk models were developed using the two approaches to variable selection, backward stepwise and the LASSO. Each approach was applied in the RUS sample ($n=1,755$) and in the combined sample ($n=2,831$). Table 5.11 displays the results from the four models developed in the RUS sample. Table 5.12 displays the results from the four models developed in the full sample.

Table 5.11: Selected model coefficients from the Fractional Polynomial and LASSO approaches in the RUS sample ($n=1,755$).

Candidate	MFP	MFP_NOAGE	LASSO	LASSO_NOAGE
t_prev_arrest			0.005	
pf_conviction				
n_probations				
n_fta	0.099	0.115	0.057	0.069
n_jails	0.161		0.066	0.005
skill	0.525	0.614	0.285	0.356
job_last_year				
havemp_r			0.167	0.164
l_drift			0.022	0.029
hasadd_r				
yrs_address			0.014	0.016
res_moves				
res_phone				
a_current			0.040	
d_current	0.258	0.276	0.168	0.184
ever_rx_a	0.457	0.342	0.259	0.196

Table 5.11: (continued)

Candidate	MFP	MFP_NOAGE	LASSO	LASSO_NOAGE
ever_rx_d				
lfam				-0.008
lfriends				
age_first_arrest		-0.024		-0.009
larceny	0.614	0.658	0.525	0.545
curr_fta				
n_pending	0.351	0.349	0.259	0.278
num_charges				
n_rec_prob				
has_alias				
gang_obs				
highgrade				
agetrncten	4.529		2.750	
anyarrest_onbail				
famcontact	1.026	0.929	0.818	0.756
notlocalyr	0.755	0.752	0.566	0.556

Table 5.12: Selected model coefficients from the Fractional Polynomial and LASSO approaches in the full sample ($n=2,831$).

Candidate	MFP	MFP_NOAGE	LASSO	LASSO_NOAGE
t_prev_arrest			0.004	
pf_conviction				
n_probations	-0.150	-0.131	-0.043	-0.036
n_fta	0.117	0.100	0.066	0.073
n_jails	0.202	0.114	0.100	0.055
skill		0.301	0.049	0.115
job_last_year				
havemp_r	0.248		0.146	0.138
ldrift	0.379		0.191	0.165
hasadd_r				
yrs_address			0.009	0.014
res_moves		0.120	0.047	0.056
res_phone				
a_current			0.003	
d_current	0.232	0.228	0.179	0.180

Table 5.12: (continued)

Candidate	MFP	MFP_NOAGE	LASSO	LASSO_NOAGE
ever_rx_a	0.307	0.243	0.130	0.087
ever_rx_d				
l_fam		-0.249		-0.027
l_friends				
age_first_arrest				-0.005
larceny	0.608	0.687	0.557	0.585
curr_fta				
n_pending	0.339	0.327	0.208	0.214
num_charges	-0.146		-0.061	-0.046
n_rec_prob				
has_alias				
gang_obs				
highgrade				
agetrncten	3.591		2.115	
anyarrest_onbail			0.150	0.164
famcontact	0.795	0.822	0.665	0.635
notlocalyr	0.529	0.484	0.377	0.372

Model Stability

When variable selection methods like MFP are used to develop a predictive model, it is important to assess the stability of the selection process. Model stability can be evaluated by calculating the bootstrap inclusion fraction (BIF) for each variable. The BIF is calculated by applying the variable selection approach in a large number ($n=200$) of random samples with replacement drawn from the original study sample and counting the fraction of times that each variable is selected for inclusion in the model. Table 5.13 shows the fraction of the bootstrap resamples in which each of the candidate variables was selected for inclusion in the model in the full sample. The top 12 variables have BIFs near 50% or higher. The top five variables were selected in almost every bootstrap replication that was conducted. For example, in the 200 bootstrap resamples and refitting of the MFP model, fam_freq was selected 199 times (99%). Note that two of the variables that we dropped (num_charges and n_probations) were selected in about 70% of the bootstrap resamples. These variables only emerge as important in the combined sample. But they have high BIFs in the LASSO approach as well (see next). However, these variables are less important in the LASSO models because the LASSO uses a different

optimization algorithm and deals with collinearity in a unique way.

Table 5.13: Bootstrap inclusion factors for MFP backward elimination approach applied to the full sample.

Candidate	Times Selected	Fraction Selected
larceny	200	100.0
agetrncten	199	99.5
fam_freq	198	99.0
notlocalyr	174	87.0
n_jails	171	85.5
n_probations	146	73.0
num_charges	134	67.0
n_pending	110	55.0
havemp_r	108	54.0
n_fta	105	52.5
ever_rx_a	103	51.5
d_current	95	47.5
anyarrest_bail	84	42.0
a_current	77	38.5
l_fam	74	37.0
l_drift	69	34.5
age_first_arrest	57	28.5
res_moves	49	24.5
t_prev_arrest	48	24.0
yrs_address	45	22.5
highgrade	43	21.5
skill	43	21.5
pf_conviction	43	21.5
hasadd_r	38	19.0
ever_rx_d	29	14.5
has_alias	28	14.0
job_last_year	26	13.0
n_rec_prob	25	12.5
l_friends	24	12.0
curr_fta	20	10.0
res_phone	16	8.0
gang_obs	13	6.5

Table 5.14 shows the BIF results for the LASSO approach in the full sample. The LASSO also selects num_charges and n_probations in a high percent of bootstrap resamples. The LASSO approach selects a broader range of variables in a large percent of the bootstrap resamples. A larger fraction of variables have BIFs above 50% in comparison with the MFP approach. However, many of the variables with high BIFs also have substantial shrinkage applied and coefficients near zero (results not shown). The bootstrap results for the LASSO approach indicate good stability as well.

Table 5.14: Bootstrap inclusion factors for the LASSO approach applied to the full sample.

Candidate	Times Selected	Fraction Selected
larceny	200	100.0
agetrncten	199	99.5
famcontact	196	98.0
notlocalyr	192	96.0
n_pending	187	93.5
n_jails	183	91.5
n_fta	177	88.5
d_current	173	86.5
havemp_r	163	81.5
num_charges	161	80.5
n_probations	153	76.5
anyarrest_onbail	146	73.0
res_moves	138	69.0
ever_rx_a	132	66.0
Ldrift	123	61.5
skill	119	59.5
a_current	110	55.0
t_prev_arrest	109	54.5
hasadd_r	104	52.0
yrs_address	97	48.5
pf_conviction	88	44.0
has_alias	88	44.0
Lfam	84	42.0
highgrade	76	38.0
n_rec_prob	73	36.5
ever_rx_d	73	36.5
curr_fta	69	34.5
job_last_year	67	33.5
Lfriends	60	30.0

Table 5.14: (continued)

Candidate	Times Selected	Fraction Selected
res_phone	51	25.5
gang_obs	51	25.5
age_first_arrest	42	21.0

Choosing a Model and Estimating Predictive Accuracy

We evaluated the predictive accuracy of the *Pretrial Release Risk Scales* using Receiver Operating Characteristics (ROC) methods. We estimated the area under the ROC curve (AUC) in the training data and in bootstrap samples to obtain estimates that compensate for over-optimistic results obtained in the training data. The results for the models developed in the RUS training data are shown in 5.15 and the results for models developed in the full sample training data are shown in 5.16.

Table 5.15: Survival Model Estimates of the area under the receiver operating characteristic curve for each risk scale model in the RUS sample.

Model	Apparent AUC	Adjusted AUC
MFP	.726	.698
LASSO	.728	.702
MFP w/o age	.704	.678
LASSO w/o age	.706	.683

Note. For Cox models the concordance index is estimated.

Table 5.16: Survival Model Estimates of the area under the receiver operating characteristic curve for each risk scale model in the full sample.

Model	Apparent AUC	Adjusted AUC
MFP	.714	.681
LASSO	.715	.692
MFP w/o age	.694	.667
LASSO w/o age	.698	.677

Note. For Cox models the concordance index is estimated.

The LASSO approach produced the best models in terms of accuracy and interpretability. The models developed in the full sample have slightly lower accuracy, but they should be more transportable than the models developed in the RUS sample. At least the full sample will resemble a broader range of pretrial settings compared with the RUS sample. Two pretrial risk scores were derived using the regression equations from the LASSO models developed in the full sample. For the two LASSO models developed in the full sample, three variables with coefficients near zero (*job_last_year*, *yrs_address*, *t_prev_arrest*) and three variables with small negative coefficients (*n_probations*, *num_charges*, *Lfam*) were removed and the model was refitted before the final equation was derived. There were significant adjustments to some of the remaining covariables, and the AUC dropped slightly (see equations below).

The default model I includes age and is calculated using the following equation:

The following transformation and recodes were applied:

- $\text{agetrncten} = (\text{age}/10)^{-2}$ - this transform changes the direction of the effect of age
- *famcontact* = recode of *fam_freq*: No family=2; Never=2; Less than 1x/mo.=2; 1x/wk=1; Daily=1
- *anyarrest_onbail* = recode of *n_arrest_on_bail*: None= 1; 1= 2; 2= 2; 3 or more = 2
- *notlocalyr* = recode of *mnth_local*: 0-2 mos. = 2; 3-5 mos. = 2; 6-11 mos.= 2; 1+ yrs. = 1

Pretrial Release Risk Scale I = $(\text{agetrncten} * 1.978) + (\text{larceny} * 0.578) + (\text{n_pending} * 0.203) + (\text{anyarrest_onbail} * 0.157) + (\text{n_fta} * 0.068) + (\text{n_jails} * 0.095) + (\text{famcontact} * 0.641) + (\text{l_drift} * 0.158) + (\text{notlocalyr} * 0.379) + (\text{res_moves} * 0.054) + (\text{d_current} * 0.193) + (\text{ever_rx_a} * 0.113) + (\text{skill} * 0.064) + (\text{havemp_r} * 0.144)$

The apparent AUC for this reduced model is .711 in the full sample. The over-optimism adjusted AUC is .688.

An alternative model II excludes age and is calculated using the following equation:

Pretrial Release Risk Scale II = $(\text{larceny} * 0.604) + (\text{n_pending} * 0.213) + (\text{anyarrest_onbail} * 0.161) + (\text{n_fta} * 0.07) + (\text{n_jails} * 0.045) + (\text{famcontact} * 0.593) + (\text{l_drift} * 0.119) + (\text{notlocalyr} * 0.374) + (\text{res_moves} * 0.068) + (\text{d_current} * 0.193) + (\text{ever_rx_a} * 0.069) + (\text{skill} * 0.132) + (\text{havemp_r} * 0.136)$

The apparent AUC for this reduced model that excludes age is .694 in the full sample. The over-optimism adjusted AUC is .673.

Failure Probabilities

The risk scores were transformed into deciles scores. The full sample was used as the reference category to cut both scores. The observed and fitted failure probabilities across the deciles of the risk scores were estimated and evaluated. Table 5.17 shows the distribution of cases across the deciles of the Pretrial Release Risk Scale I, including the number that fail within each decile, the observed proportion failing in each decile level, and the crude cumulative incidence. The crude cumulative incidence is the probability of FTA or felony arrest adjusted for losses from other causes (death, case dismissal, technical violation, and success). The observed probability and cumulative incidence estimates are quite similar. There is a consistent trend of increasing probability of failure as one moves up the deciles of the Pretrial Release Risk Scale, with the exception of a flattening near D7 to D8. This reflects the shape of the distribution of the risk score, which displays kurtosis (peaked distribution). This is also an indication of lack of sharpness - a large proportion of the predicted probabilities from the risk score fall in the medium range.

The observed probability of FTA or Felony Arrest in the first, fifth and tenth deciles of the Pretrial Release Risk Scale I is 0.03, 0.1, and 0.23, respectively.

If we cut the deciles at D4 and D9 (1-3=low, 4-8=medium, 9-10=high), there are 548 (19%) individuals classified as high risk on this scale.

Table 5.17 shows the proportion of cases failing within the deciles of the Pretrial Release Risk Scale II. The results are quite similar to the those described above for model II.

Figure 5.2 shows a plot of the fitted failure probabilities (1 - KM) of any arrest within the three levels of the Pretrial Release Risk Scale I. The plot shows that we have good separation in fitted probabilities between the Recidivism Risk levels if we use cut points at D4 and D9. Notice that we have good precision in the estimates over the follow-up (up to 240 days), as evidenced by the narrow width of the 95% error bars. The table on the top of the graph shows the size of the riskset over time in the levels of the Recidivism Risk scale. The risk set gets smaller over time as cases are lost to failure (FTA or New Felony Arrest) and censoring. In the cause-specific models that we fitted, cases are censored (removed from the risk set) if they experience a competing event (success, technical violation, case dismissal, or case reduced and transferred).

Figure 5.3 shows a plot of the fitted failure probabilities (1 - KM) of FTA or Felony within the three levels of the Pretrial Release Risk Scale II, that excludes age. The Pretrial Release Risk Scale II is also cut at D4 and D9. This results in 50% of the cases falling into the medium risk level. When the Pretrial Release Risk Scale is deployed in a different jurisdiction, new deciles and cutting points should be set and tested.

Table 5.17: Number of cases, proportion with FTA or felony arrest, and crude cumulative incidence across the Pretrial Release Risk Scale I deciles

Pretrial Risk Scale Decile	Number in Level	Number Failing	Proportion Failing	Cumulative Incidence
D1	294	7	0.02	0.03
D2	303	12	0.04	0.04
D3	283	15	0.05	0.04
D4	268	25	0.09	0.10
D5	295	30	0.10	0.10
D6	289	31	0.11	0.11
D7	267	19	0.07	0.08
D8	284	30	0.11	0.10
D9	281	55	0.20	0.20
D10	267	62	0.23	0.24

Note. Crude cumulative incidence estimate at 240 days from release ($n=2,831$).

Table 5.18: Number of cases, proportion with FTA or felony arrest, and crude cumulative incidence across the Pretrial Release Risk Scale II deciles

Pretrial Risk Scale Decile	Number in Level	Number Failing	Proportion Failing	Cumulative Incidence
D1	301	10	0.03	0.04
D2	288	15	0.05	0.06
D3	304	20	0.07	0.06
D4	264	22	0.08	0.08
D5	286	28	0.10	0.10
D6	284	28	0.10	0.09
D7	277	21	0.08	0.08
D8	273	28	0.10	0.11
D9	288	52	0.18	0.18
D10	266	62	0.23	0.24

Note. Crude cumulative incidence estimate at 240 days from release ($n=2,831$).

Note that the jump at 15 days in the survival curves for both risk models reflects the FTA failures at the first examination hearing following arrest. There were 34 defendants in the high-risk group that failed on day 15, compared with 6 in the low risk group.

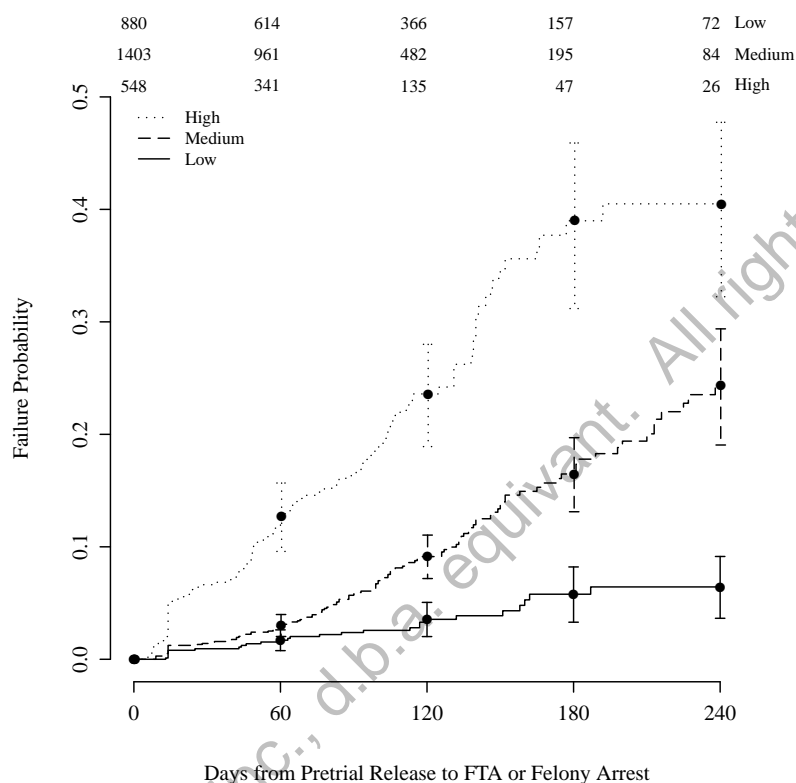


Figure 5.2: Plot of the naive failure probability (1-KM) of FTA or Felony Arrest within the levels of the Pretrial Release Risk Scale I with cuts at D4 and D9.

In typical survival data without competing events, the Kaplan-Meier statistic is used as an estimate of survival or failure probability (1 - KM) at different time points. However, in the context of competing risks, 1 - KM is not a proper or interpretable failure probability. Usually 1 - KM will be biased upward. One recommended approach is to plot the predicted values from a regression of the cumulative incidences on the covariates (i.e., the levels of the Pretrial Release Risk scale) (Fine & Gray, 1999). Figure 5.4 shows such a plot of the fitted failure probabilities of FTA or felony arrest within the three levels of the Pretrial Release Risk Scale I. A comparison of Figure 5.4 with Figure 5.2 reveals the classic over-estimate of the failure probability obtained with the naive Kaplan-Meier statistic from the cause specific model.

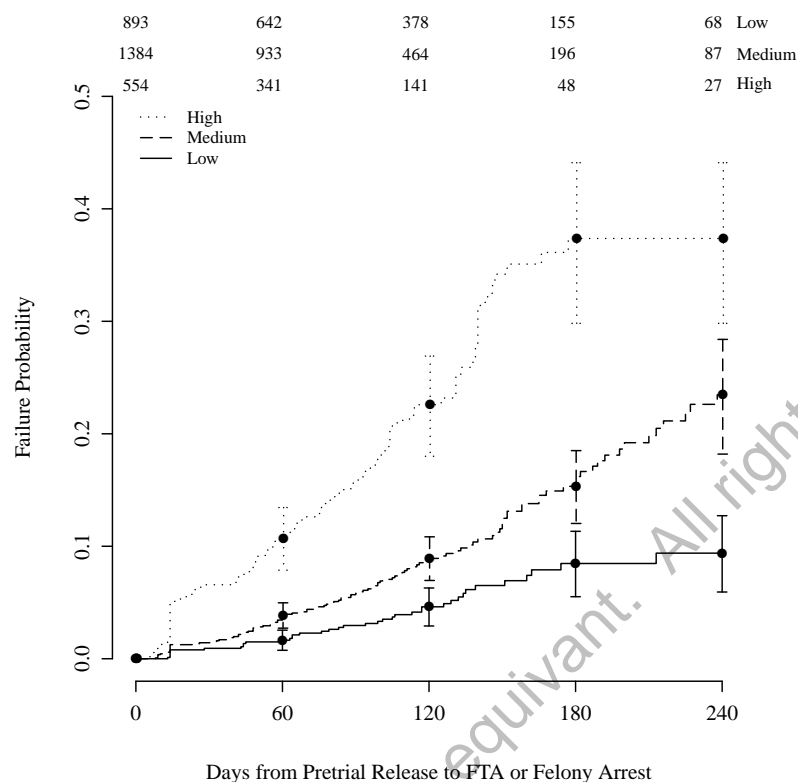


Figure 5.3: Plot of the naive failure probability (1-KM) of FTA or Felony Arrest within the levels of the Pretrial Release Risk Scale II with cuts at D4 and D9.

Figure 5.5 shows a plot of the fitted failure probabilities of FTA or felony arrest within the three levels of the Pretrial Release Risk Scale II. The plot shows that there is only modest separation in fitted probabilities between the low and medium levels using the default cut points at D4 and D9. Again a comparison with the naive KM-1 failure probability in Figure 5.3 shows the typical upward bias. The naive KM-1 failure probability is almost double the crude cumulative incidence function that accounts for the competing events.

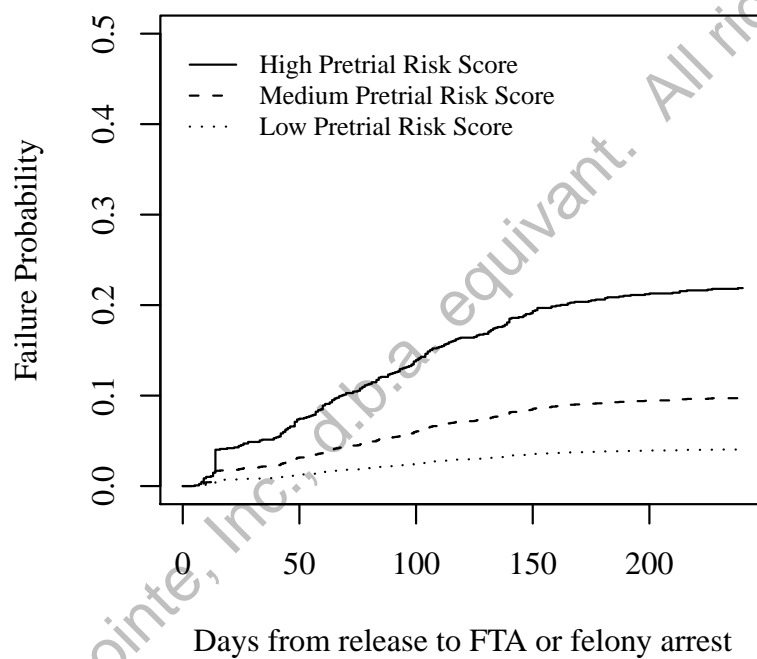


Figure 5.4: Predicted cumulative incidence of any arrest within the levels of the Pretrial Release Risk Scale I with cuts at D4 and D9.

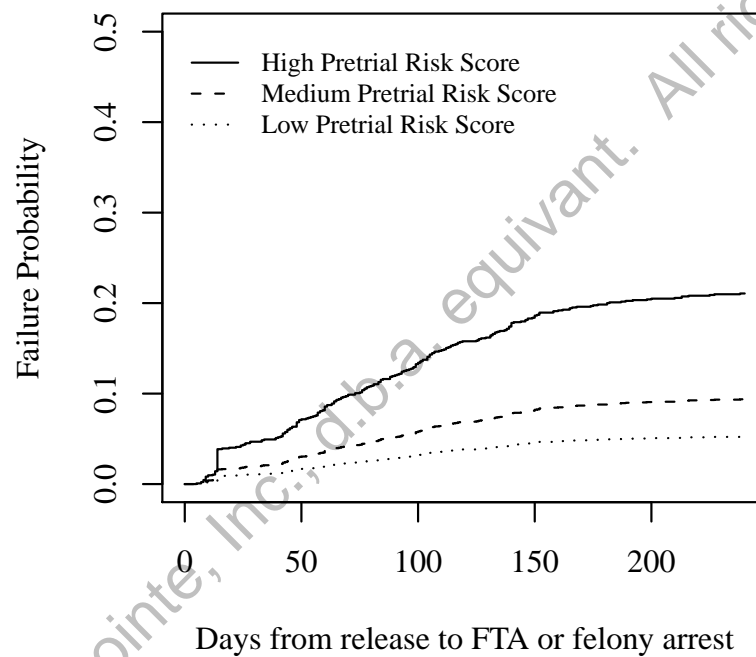


Figure 5.5: Predicted cumulative incidence of any arrest within the levels of the Pretrial Release Risk Scale II with cuts at D4 and D9.

References

- Cohen, T. H., & Reaves, B. A. (2004). *Pretrial release of felony defendants in state courts* (Tech. Rep.). Washington:D.C.: Bureau of Justice Statistics.
- Cuvelier, S. J., & Potts, D. W. (1993). *Bail classification profile project: Final report* (Tech. Rep.). Alexandria, Virginia: State Justice Institute.
- Division of Probation and Correctional Alternatives. (2007). *New York State pretrial release services standards*.
- Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446), 496–509.
- Goeman, J. J. (2010). L1 penalized estimation in the Cox proportional hazards model. *Biometrical Journal*, 52(1), 70-84.
- Goldkamp, J., & Gottfredson, M. (1985). *Guidelines for bail: The Philadelphia experiment*. University Press.
- Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L., & Rosati, R. A. (1982). Evaluating the yield of medical tests. *Journal of the American Medical Association*, 247, 2543-2546.
- Harrell, F. E., Jr. (2001). *Regression modeling strategies*. New York: Springer.
- Lowenkamp, C. T., Lemke, R., & Latessa, E. (2008). The development and validation of a pretrial screening tool. *Federal Probation*, 72, 2-8.
- Lowenkamp, C. T., & Whetzel, J. (2009). The development of an actuarial risk assessment instrument for U.S. Pretrial Services. *Federal Probation*, 73, 33-36.
- National Association of Pretrial Services Agencies. (2004). *National Association of Pretrial Services Agencies standards on pretrial release (3rd edition)* (Tech. Rep.). National Association of Pretrial Services Agencies.
- Pintilie, M. (2007). Analysing and interpreting competing risk data. *Statistics in Medicine*, 26, 1360 - 1367.
- Putter, H., Fiocco, M., & Geskus, R. B. (2007). Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine*, 26, 2389–2430.

- Rhodes, W., Hyatt, R., & Scheiman, P. (1996). Predicting pretrial misconduct with drug tests of arrestees: Evidence from eight settings. *Journal of Quantitative Criminology*, 12, 315-348.
- Sauerbrei, W., & Royston, P. (1999). Building multivariable prognostic and diagnostic models: Transformations of the predictors by using fractional polynomials. *Journal of the Royal Statistical Society, Series A*, 162, 71-94.
- Siddiqui, Q. (2005). *An evaluation of the new pretrial release recommendation system in New York City: Phase II of the post-implementation research* (Tech. Rep.). New York: New York City Criminal Justice Agency.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, 58, 267-288.
- VanNostrand, M. (2003). *Assessing risk among pretrial defendants in Virginia: The Virginia pretrial risk assessment instrument* (Tech. Rep.). Virginia Department of Criminal Justice Services.
- VanNostrand, M. (2007). *Legal and evidence-based practices: Application of legal principles, laws, and research to the field of pretrial services* (Tech. Rep.). National Institute of Corrections and Crime and Justice Institute.
- VanNostrand, M., & Keebler, G. (2009). *Pretrial risk assessment in the Federal Court: For the purpose of expanding the use of alternatives to detention* (Tech. Rep.). Washington, D.C.: Department of Justice.
- Visher, C. A., & Linster, R. L. (1990). A survival model of pretrial failure. *Journal of Quantitative Criminology*, 6, 153-184.
- Winterfield, L., Coggeshall, M., & Harrell, A. (2003). *Empirically-based risk assessment instrument: Final report* (Tech. Rep.). Urban Institute Justice Policy Center.