Pretrial Failure
Among New York City Defendants
By Qudsia Siddiqi, Ph.D.

In New York State, the Criminal Procedure Law (CPL) prohibits the explicit consideration of public safety in the setting of pretrial release/detention conditions. Therefore, the pretrial release recommendations made by CJA are based solely on a defendant’s risk of flight.

Our previous research showed lower failure-to-appear (FTA) rates among defendants considered by the Agency to be at low risk of flight, compared to defendants at moderate or high risk of flight (Research Brief no. 13). However, the question arises whether those defendants at low risk of failure to appear would also be at low risk in terms of the safety of the community.

In response to this question, we conducted research that focused on both types of risk among New York City defendants: public safety as well as flight. We used pretrial re-arrest as a measure of public safety. Pretrial failure is the term used to denote the combined types of failure.

The research focused on the following questions:

- What factors predict pretrial failure?
- How could these factors be incorporated in a risk classification system that predicts pretrial failure?
- How does this failure risk classification system compare with CJA’s ROR recommendation system?

The major findings from the study are presented in this Brief.
Plan of Analysis

First, multiple logistic regression models were developed to identify statistically significant predictors of pretrial failure. Multiple logistic regression analysis is a statistical technique that is used to test the individual effects of a number of independent variables on a dichotomous dependent variable, while controlling for the other variables in the model. The dependent variable in this study was pretrial failure, defined as a failure to appear or a re-arrest prior to disposition of the case.

Statistics produced by the logistic regression analysis were used as the basis for assigning negative or positive points to each item that significantly affected the likelihood of pretrial failure. A score for each defendant was produced by tallying the points for all items on the scale. Based on these scores, a classification system was then devised to predict each defendant’s level of risk for pretrial failure.

Finally, the resulting failure risk classification system was compared with the recommendation system that is currently in use by CJA to predict risk of flight.

Selected Characteristics of Defendants in the Sample

As shown in Figure 1 (top row), the majority of defendants were male (84%), and slightly less than one half were black (47%). The median age was 30 (not shown). Defendants initially arrested on a felony charge comprised about half the sample (51%).

An attempt is made by CJA to verify community ties information collected at the pre-arraignment interview by calling a contact number provided by the defendant. The interview items in Figure 1 (middle row) are categorized as verified (YV=yes verified; NV=no verified); not verified (Y=yes; N=no); or the defendant and the contact person gave conflicting information (UC=unresolved conflict). Most defendants reported living in the NYC area (93%, combining Y and YV); a majority had a telephone (75%, combining Y and YV); and slightly over half reported being employed, in school, or in a training program full time (54%, combining Y and YV).

Criminal history data (bottom row) show that 29% of defendants had a prior misdemeanor conviction, 24% had an open case at the time of arrest, and 26% had failed to appear for a court date on at least one previous occasion.

The Sample

Data for the present analysis were drawn from a cohort of prosecuted arrests made between January 1 and March 31, 2001, in New York City. Defendants who had been issued a desk appearance ticket were excluded from the study sample. Only the first arrest during the sampling period for each defendant was included. After these initial exclusions, the number of cases in the dataset was 67,848. Cases were tracked in Criminal Court until November 30, 2001, and in Supreme Court until January 31, 2002.

Our analysis was limited to adult defendants whose cases were not completed at arraignment and who were released prior to the disposition of the case in Criminal or Supreme Court. Defendants who were held in detention for the duration of their cases were excluded because they were never at risk for FTA or re-arrest. Juveniles and defendants who were not interviewed by CJA were also excluded. The final sample used in these analyses consisted of 26,820 adult defendants who were at risk for pretrial failure.
Predicting pretrial failure

Figure 2 presents the distribution of pretrial FTA and/or re-arrests among defendants in the sample. The overall pretrial failure rate was 28%, distributed as follows: 10% failed to appear for a scheduled date in either Criminal or Supreme Court prior to disposition; a slightly higher proportion (12%) were re-arrested pretrial; and 6% failed in both ways (re-arrest and FTA).

The first step in developing a risk assessment instrument for pretrial failure was to identify significant predictors of pretrial failure. To achieve that objective, logistic regression models were developed (see previous page). For the present analysis, the effect of a factor was considered statistically significant if it had a 5% or less probability of having occurred by chance. The results from the final model are presented in Figure 3.

Figure 3
Multivariate Logistic Regression Model Predicting Pretrial Failure
(N=25,399)

The lines in Figure 3 represent logit coefficients, which reflect the impact of the predictor variable on the dependent variable, while controlling for other variables in the model. The positive or negative values indicate that the predictor variable increases or decreases the likelihood of the outcome. In our model, the coefficients reflect the change in the likelihood of pretrial failure associated with a change in the predictor variable.

The length of the line represents the size of the coefficient: the larger the coefficient (and the longer the line), the greater the importance of the item in predicting pretrial failure.

Lines above the base depict positive coefficients, indicating that the characteristic was associated with a greater likelihood of failure. Lines extending below the base depict negative coefficients, indicating reduced likelihood of failure.

The coefficients for the criminal history items indicate that these prior events are associated with a greater likelihood of failure, compared to the absence of the event. The coefficients for the community ties items indicate the greater or lesser likelihood of failure associated with each response, compared to the average failure rate for the sample.

All of the items in the model were statistically significant, with the exception of N and NV responses to the NYC area address question and the UC response to the telephone question. Bars depicting coefficients that were not significant (n.s.) are gray rather than blue or black.

Having a prior FTA was the factor with the strongest effect on likelihood of failure (.636). Additional criminal history factors that increased a defendant’s likelihood of failure were having another case open at the time of arrest (.347) and having a prior misdemeanor conviction (.147).

Three community-ties variables also affected likelihood of failure. Defendants who reported a New York City area residence—especially those for whom the information was verified—were less likely to fail (YV, -.256). (“Yes” responses to community-ties items are shown in blue, other responses in black.)

Similar results were obtained for having a telephone and a full-time activity (employment, school or a training program). The presence of either a telephone or a full-time activity reduced the likelihood of failure, and the absence of either increased the likelihood of failure.
Developing a Failure Risk Classification System

Constructing a Point Scale for Pretrial Failure

The regression coefficients shown in Figure 3 were used to develop a point scale for targeting defendants at risk of pretrial FTA or re-arrest. Each statistically significant coefficient was divided by a constant (.15) and rounded to the nearest whole number to produce a point scale value. The signs were reversed in constructing the point scale because a negative coefficient means less likelihood of failure, and the point scale is designed to assign higher scores to defendants less likely to fail. The coefficients that were statistically insignificant were assigned a value of zero. Also, due to the small proportion of defendants with an unresolved conflict on the NYC area address item, a value of zero was assigned to that category.

For example, the logit coefficient for a prior FTA was .636, which divided by .15 results in a value of 4 points for this item (rounded from 4.24). Reversing the positive sign of the coefficient (to –4) results in a loss of 4 points for a defendant with a prior FTA. This item contributed the most to a defendant’s total score because the regression analysis showed that it had the greatest impact on likelihood of pretrial failure.

Figure 4 presents the 6-item point scale for predicting pretrial failure (blue), overlaid on the point scale currently used in the recommendation system for release on recognizance (gray). The scales are very similar, but one item on each scale is not used in the other. Expecting someone at arraignment is useful in predicting FTA (for the ROR recommendation), but not combined failure. A prior misdemeanor conviction is useful in predicting combined failure, but not FTA alone.

Points were tallied to produce a total score for each defendant, ranging from –9 to +10. Figure 5 shows the failure rate associated with each score. As expected, defendants with the lowest score had the highest failure rate (60%) and those with the highest score had the lowest failure rate (14%).

The distribution of scores is also shown in Figure 5 (bottom line). The most common scores were +7 through +10, with 10% or more of the defendants in the sample qualifying for each of the four highest scores. The least common score was –9 (fewer than 1% would have received a score that low).
Classification of Defendants At Risk For Pretrial Failure

The point scale predicting pretrial FTA and/or re-arrest was used to categorize defendants by their level of risk for failure. The same cutoff points used in the ROR recommendation system were applied to create three categories of risk: **low** (7 to 10 points), **moderate** (3 to 6 points), and **high** (–9 to 2 points). (The ROR recommendation system has a range from –12 to +12). Included with the high-risk category were defendants who are not recommended for ROR or for whom no release recommendation is made for policy reasons (outstanding bench warrant, homicide or bail jumping charge, conflicting residence information, unavailable criminal history information, or incomplete interview). Figure 6 presents the resulting distribution of failure risk categories. Under this classification system, 48% of defendants in the sample would be considered low risk, 19% moderate risk, and 33% high risk.

Comparing the Failure Risk Classification System With the ROR Recommendation System

We compared the failure risk classification system with the ROR recommendation system in terms of the proportion of defendants in each risk category and their corresponding failure rates. Figure 7 shows that 42% of defendants in the sample would be categorized by the ROR recommendation system as low risk, fewer than under the failure risk classification system (48%, Figure 6). The ROR recommendation system would classify about the same proportion as moderate risk (20% compared to 19%), and a larger proportion as high risk (38% compared to 33%).

Comparisons between the two systems in failure rates for each risk category are presented in Figure 8. The failure rate for defendants classified as low risk under the failure risk classification system was higher than under the ROR recommendation system (18% compared to 9%), but the two rates are not directly comparable. The combined failure rate is higher because it counts re-arrest and FTA as failure, whereas the ROR recommendation system counts only FTA.

A better comparison lies in the difference between the failure rate of the low-risk group and the average for the sample as a whole (base rate) under each system. The base rates are shown as gray backdrops in Figure 8. The low-risk group under the failure risk classification system had a failure rate that was 10 percentage points lower than the average for the sample as a whole: 18% compared to the base rate of 28%. By comparison, the low-risk group under the ROR recommendation system had a failure rate only 7 points below the base rate: 9% compared to 16%. We conclude that the failure risk classification system would be just as effective as the ROR recommendation system in identifying low-risk defendants.

Moreover, high-risk defendants under the failure risk classification system had a failure rate that was 12 percentage points higher than the base failure rate (40% versus 28%), compared to an 8-point difference under the ROR recommendation system.
Finally, we compared the two risk assessment systems on all the measures of failure: FTA, re-arrest, and combined failure. Earlier comparisons assessed each system only in terms of the risk it was designed to predict (combined failure for the failure risk classification system and FTA for the ROR recommendation system). Now we compare the ability of the two systems to predict each measure.

Figure 9 shows that, even though the failure risk classification system was not designed to predict FTA alone, the FTA rates for low-risk (9%) and high-risk (24%) defendants were the same under both systems.

Figure 10 shows that, even though the ROR recommendation system was not designed to predict re-arrest, it predicted re-arrest nearly as well as the failure risk classification system. Re-arrest rates by category were very similar, regardless of which system was used to classify defendants: 11% for low-risk defendants under either system, and from 24% (ROR) to 26% (failure risk) for high-risk defendants.

Figure 11 shows that the combined failure rates were also the same under the two systems, when the rates for low-risk (18%) and high-risk (40%) defendants are compared.

The differences, although small, were mostly among the moderate risk categories. Defendants classified as moderate risks under the ROR recommendation system had slightly lower rates on all three measures than defendants classified as moderate risks under the failure risk classification system.

Comparing the Two Systems on Predicting Three Measures of Failure

The new classification system presented in this report categorized defendants by their risk of pretrial failure. Relative to the ROR recommendation system currently in use by CJA, the failure risk classification system would categorize more defendants as low risk. In addition, the failure rate for low-risk defendants would represent a greater reduction from the average failure rate than the parallel reduction in FTA for low-risk defendants under the CJA recommendation system. It would classify the same proportion of defendants as moderate risk, and this group would have a slightly higher failure rate than the sample’s average. Finally, it would classify fewer defendants as high risk, and the failure rate for this group would be considerably higher than the average.

The increase in the proportion of low-risk defendants under the failure risk classification system was mainly due to a shift in points assigned to some of the factors that predict both FTA and combined failure. These include reporting a New York City area address and having a history of prior FTA. Both of these factors weighed less heavily in the failure risk classification system than in the ROR recommendation system.

The substitution of one item in the ROR recommendation system (expecting someone at arraignment) with a different item (prior misdemeanor conviction) also helps to account for the increase in the proportion of low-risk defendants under the combined failure system. Under the combined system, defendants with a prior misdemeanor conviction would lose one point, whereas defendants with no such conviction would earn one point. Under the ROR recommendation system, one point would be deducted if the defendant did not expect a relative or friend at arraignment, whereas a point would be added for expecting someone at arraignment.

One might argue that since the two variables were given the same weight in the separate systems, the substitution of one for the other would not affect a defendant’s total score. Defendants would earn or lose one
point irrespective of which variable was included in the scale. However, the proportion of defendants with no prior misdemeanor conviction (71%, Figure 1) was somewhat higher than the proportion of defendants who expected someone at arraignment (62%, not shown). Consequently, the proportion of defendants earning one point for not having a prior misdemeanor conviction would be higher under the combined system than the proportion of defendants scoring one point for expecting someone at arraignment under the ROR system.

The two systems were compared with respect to their ability to predict FTA, re-arrest, and combined failure. On all three measures, the rate for the low-risk group was exactly the same regardless of which system was used to classify the risk level.

### Implications

Although the current New York State statute does not permit the consideration of public safety in making pretrial release or detention decisions, our findings suggest that failure to appear is closely linked to pretrial re-arrest. Defendants considered low risks for FTA also have the lowest re-arrest rate. By the same token, defendants categorized as high risk for FTA also have the highest re-arrest rate. By predicting risk of FTA, the CJA recommendation system also predicts risk of re-arrest.

These findings raise questions about releasing high-risk defendants on recognizance. Currently, other release conditions, such as supervised release or release under third-party custody, are not an option in New York City. The current statute limits CJA’s ability to make recommendations that would take public safety into consideration, but conditional release options aimed at reducing a defendant’s risk of flight would also affect risk of re-arrest. We expect that a reduction in flight risk would be followed by a reduction in pretrial criminality. This suggestion would be appropriate for moderate-risk, as well as high-risk, defendants.

Should legislative changes be made in New York that would allow consideration of public safety in making pretrial release decisions, some further work will need to be done to refine the classification procedures. For example, the model used to develop the pretrial failure risk classification system did not control for the effect of the arrest charge. In research a few years ago, we found that arrest charge severity and offense type were significant in predicting FTA, but this item was not included in the point scale for the ROR recommendation system, partly because it implied that positive points should be awarded for certain offenses, such as gambling and driving under the influence of alcohol or drugs. At the suggestion of Criminal Court judges, and due to the operational difficulties in awarding positive or negative points for different offense types and severities, this variable was excluded from the ROR recommendation system. Variables reflecting charge type and severity were also excluded from the present analysis of combined failure, in order to be consistent with the current recommendation policy. The decision regarding the charge variables will need to be revisited in the event that public safety considerations are ever allowed in making release and detention decisions in New York.

In addition, the cutoff scores used here for grouping defendants into low-, moderate-, and high-risk groups are not necessarily optimal. When developing the failure risk classification system, we used the same cutoff points as are used in CJA’s ROR recommendation system. This was done for research purposes, so that we could compare the performance of the two systems. Should the statute allow consideration of public safety in making pretrial release decisions, those cutoff points will need to be redefined, based upon a determination by criminal justice professionals of an acceptable level of failure for the low-risk group.

Finally, the findings presented here should be validated at such time as any change occurs in the permissible considerations for pretrial release decisions. The findings presented in this report are based on data that were collected prior to the implementation of the current ROR recommendation system in 2003 and should be validated on post-implementation data.
Research Brief from

No. 19 (January 2009)
Pretrial Failure Among New York City Defendants

Forthcoming:
No. 20 (May 2009)  The Manhattan Day Custody Program
by Freda F. Solomon, Ph.D.

Most recently published in this series:
No. 18 (September 2008)  Bail, Detention, and Felony Case Outcomes (Phillips)
No. 17 (May 2008)  Pretrial Misconduct Among Domestic Violence Defendants (Peterson)
No. 16 (January 2008)  Predicting The Likelihood of Pretrial Re-Arrest For Violent Felony Offenses (Siddiqi)
No. 15 (September 2007)  The Risk of Re-Arrest For Serious Juvenile Offenders (Gewirtz)
No. 12 (September 2006)  Pretrial Outcomes For Domestic Violence Defendants (Peterson)
No. 11 (April 2006)  New York’s Gun Court Initiative: A Pilot Program Study (Solomon)
No. 10 (December 2005)  Assessing the Impact of Differing Models of Youth Crime Prosecution (Gewirtz)
No. 9 (August 2005)  Prosecutors’ Bail Requests and the CJA Release Recommendation (Phillips)

www.nycja.org/research/research.htm

Research Brief from

No. 19 (January 2009)
Pretrial Failure Among New York City Defendants

The New York City Criminal Justice Agency, Inc.
52 Duane Street
New York, NY 10007