

Colorado Pretrial Assessment Tool (CPAT)

FAQs

Prepared by the Pretrial Justice Institute

Updated: October, 2012

What else besides the CPAT itself does a jurisdiction need to do to implement the CPAT and maximize its value?

To help with implementation, an administration and scoring manual is being prepared for anyone who wishes to use the CPAT. This manual will help standardize the CPAT's use across jurisdictions and across staff within a single jurisdiction, as well as help assure that the CPAT is administered and scored with accuracy.

To maximize the CPAT's value, it is important to view the CPAT itself as just the first part of improving the pretrial assessment, recommendation, and supervision process for the local pretrial process and program. It will also be necessary to develop a matrix/praxis (for example, see Mesa County's matrix, which has been reviewed by and received input from national-level pretrial experts). The matrix is necessary for a variety of reasons, including measuring necessary outcomes for your system and program (e.g., public safety rates, court appearance rates, release rates) for determining if the CPAT is, even unintentionally, being used to over-incarcerate lower risk defendants, to over-supervise lower or medium risk defendants, and/or in a manner biased against minorities. Furthermore, the matrix helps guide judge's decision-making on which defendants to release or detain and which bond conditions can be used to reduce released defendants' pretrial misconduct, and it helps mitigate practices that inadvertently undermine that decision-making.

What risk categories should be used when linking a defendant's CPAT score to a matrix/praxis -- the original categories in the CPAT's initial report (dated February 2012) or the revised categories issued in August 2012?

The revised categories should be used. Please refer to the separate document that describes the new categories.

Is it okay to use the CPAT to assess minority defendants' risk?

Yes. The analysts who developed the CPAT tool (JFA Institute) said that they used scientifically accepted procedures to develop the tool. These procedures include analyses to test whether the tool is biased based on demographic factors. Additionally, both the JFA Institute and a third party reviewer with expertise in pretrial in risk assessment instrument said that the CPAT does a good job of indicating the pretrial risk level of both White and minority defendants. Furthermore, the third party expert and the Pretrial Justice Institute strongly recommend that each jurisdiction perform ongoing measurement of how the CPAT is used to assure it does not contribute to decision-making biased against minorities. A good way to do this is to measure within each cell of the local jurisdiction's bond conditions matrix (i.e., a tool that cross-references CPAT risk

category with top charge and other factors that influence the most appropriate set of bond conditions to manage a defendant's pretrial risk) not only the public safety and court appearance rates, but the release (from jail) rates as well. Measuring the release rates in each cell in total and for subtypes of defendants (e.g., by race or ethnicity, or by sex), jurisdictions can determine if the local court's bond setting practices are inadvertently biased against minorities or other subgroups. If any biases are identified, PJI is committed to assisting any local jurisdiction in modifying its policies and procedures to eliminate the bias.

The current charge has often been used as an indicator of the degree of pretrial risk, either through a bond schedule or in court, but it is not included in CPAT. Is there any support at all for using top charge to measure risk in the study's results?

See first paragraph on page 20 of the full report ("...this study failed to show that the nature (e.g., person or property crime) or severity (felony, misdemeanor) of the defendant's current charge was statistically significantly related to pretrial misconduct"). Current charge did not make the cut for being related to pretrial misconduct. If there is any correlation, it's due to chance or can be accounted for by other factors. So, current charge is not as good as a combination of other factors in assessing likelihood of pretrial risk in CO at this time. The current charge factor is predictive in some other jurisdictions, and may or may not be included in the next version of the CPAT when it is revalidated in the future.

As for support for using current charge (or even a risk score from a combination of factors like on the CPAT) to determine money bond amounts on a printed schedule, that is a separate, somewhat complex issue. Several pretrial experts have presented compelling legal, rational, and statistical reasons for not doing that any more. Most all (or all) Colorado jurisdictions except Jefferson County use a printed money bond schedule based on current charge alone (and the Jeffco courts still sometimes link higher risk of pretrial misconduct to higher monetary bond amounts). This issue goes beyond the new tool itself, and discussions about it may arise in jurisdictions that implement the CPAT.

Current charge is an important consideration when judges set the conditions of bond, as is input from the prosecutor, defense attorney, victim, the defendant, and/or the defendant's family. Judges weigh all of this information when setting bond conditions. In sum, current charge type or severity does not measurably contribute to a defendant's likelihood of pretrial misconduct, but it does factor in the judge's decision about how to handle the defendant's risk and the criminal case overall.

Why are the items for the number of prior FTAs and number of adult felony convictions not on the CPAT?

The analysts who performed the statistical analyses provided the following information: The current risk tool does a good job of identifying defendants' risk of FTA and getting a new filing, or both. The items used in the instrument are similar to ones used in other pretrial risk instruments used in other jurisdictions. Figures 1 and 2 in the report demonstrate this. On the Prior FTA matter, as noted in the report, there is a strong bivariate relationship between Prior FTA and the pretrial misconduct outcomes (FTA and new filing). But there is another item, Prior

Revocations, which is equally or even more strongly related to the pretrial misconduct outcomes. The same can be said about the item Prior Adult Felony Convictions. It too was strongly associated with FTA and new filing, but so too were other measures of prior criminal behavior that had the same or stronger associations with the outcomes of interest. The purpose of the regression model (multivariate analysis) is to identify the best and most efficient set of predictors for the tool. In a regression model it's important not to include items that are strongly correlated with one another. This problem is called "multicollinearity."¹ In our analysis, the presence of the other prior and current criminal factors, prior FTA, and prior adult felony convictions do not add to the predictive efficacy of the models. Therefore, these variables were dropped from the final model. In the final model, 7 of the final 12 items are related to prior or current criminal conduct. One could add the items Prior FTA or Prior Adult Felony Convictions, but they will not improve the overall predictive efficacy of the tool. In fact, they may diminish it slightly.

Does having 8 of the 12 items based on the defendant's self-report reduce the validity of the CPAT?

Although this is possible, it is not likely for the following reasons. First, these 8 items "beat out" many other items that were not self-reported when the best predictors of pretrial risk were identified in the statistical analyses. Second, although some defendants will give incorrect responses to self-report items either intentionally or unintentionally, this "noise" in the data was present when the items were identified in the CPAT's development, and they were still predictive (and more so than many other items). Third, judges, court clerks, and pretrial program staff from multiple jurisdictions have said that the data in the court's information system and defendant's criminal history is not always accurate, so it is not always accurate to assume that a defendant's response is false if it contradicts the data in the information system. That is, defendants' self-reported information can be more reliable than data in the information system. Finally, the recommended ongoing standardized data collection that accompanies the CPAT's use consists of data collection on the CPAT's 12 items as well as other factors that may be better predictors in the future. Thus, any of the CPAT's 12 items, including the 8 self-report items, may remain on or be replaced by other items in future updated versions of the CPAT.

Why are some factors that statute says the judge shall consider not in the CPAT?

There are a few reasons for this:

1. Many of the factors listed in CO statute, like those in many other states, were copied from other states' statutes several decades ago. At the time, the state of the science of pretrial risk assessment nationally was not as advanced as it is today. So the factors listed in statute were and are based on less scientific criteria, whereas the tool has the most optimal set of 12 factors that could be derived from the data collected in Colorado (which was more thorough than that typically collected in other jurisdictions for pretrial risk instrument development).

¹ Multicollinearity in logistic regression models is a result of strong correlations between independent variables (potential predictors). The existence of multicollinearity inflates the variances of the parameter estimates. That may result, particularly for small and moderate sample sizes, in lack of statistical significance of individual independent variables while the overall model may be strongly significant. Multicollinearity may also result in wrong signs and magnitudes of regression coefficient estimates, and consequently in incorrect conclusions about relationships between independent and dependent variables.

2. In most states, including CO, pretrial statutes contain items that go beyond the ability of statistics to measure, such as the “apparent probability of conviction and the likely sentence.” Because statutes are not just written from a basis of research and statistics alone, they will contain non-statistically-derived content.
3. As sometimes occurs, statutes are written or changed for political or other reasons that are not always based in science or research. So, it is possible that factors were added to statute over time for political or other non-scientific reasons.

So, Colorado statute tells judges what 14 factors they must consider, but it does not tell them what weight to assign each factor or how to define or measure the factors, either individually or in combination, when making their bond decisions. For the reasons above, the CPAT does not, and cannot, address every statutory factor.

In addition, statute says that judges shall consider “defendant’s employment status and history.” The CPAT research showed that this specific item was not as strongly linked to pretrial misconduct as were other items when all items were considered together. Judges statutorily still have to consider this factor, so they can choose to give that factor the weight the Colorado research says it deserves. It would not be practical to include this item on the tool and score it a zero for all defendants.

Is there any research on defendants who refuse or are uncooperative to be interviewed as an indicator of risk?

This research has not been done or published in the pretrial field. Part of the reason is that there are very few defendants who refuse compared to defendants who participate in the interview. In addition, the number of initial refusers decreases even further with skilled pretrial staff and certain court actions (e.g., the judge deferring the bond setting to the next day to give the defendant another chance to participate in the interview). The refusal rate for the CPAT study was less than 1%.

Why was “any previous criminal justice related supervision in the past 10 years” (p. 10) removed from the final model? Since it’s negatively associated with pretrial misconduct, which could mean that it actually reduces the risk, could it be included as a negative score?

There is an association between the “previous supervision” item and pretrial misconduct in the bivariate analysis, which is consistent with the research literature. However, in the subsequent multivariate analysis, it develops a negative relationship to pretrial misconduct. This is likely an artifact of how the various factors interact with each other when their unique association to pretrial misconduct is taken into account. This does not happen often but it does occur.

Of the 12 factors listed on page 12, some are listed only under “FTA,” some only under “New Filing,” and others only under “Either FTA or New Filing.” Does this argue in favor of separate scores for FTA and new filings?

Please see footnote 18 on page 13. Please also see Table 4 on page 18, where a defendant’s CPAT score is translated separately for public safety success and court appearance success.

Factors related to only one form of pretrial misconduct but not the combined outcome, or being related only to the combined outcome measure but not the individual outcomes, occurs because the relationship can change when a composite or combined outcome measure is used versus separate outcomes. This phenomenon occurs because the combined failure outcome uses a combined and larger base rate measure instead of a single and smaller base rate measure. It becomes increasingly more difficult to develop risk assessment models as the base rates of failure decrease. This is why almost all pretrial risk assessment instruments assess a combined measure of pretrial misconduct. Nonetheless, the distinction between likelihood of any given defendant's risk of failure to appear and for being charged with a new crime is important to the court, so Table 4 on page 18 separates defendants' likelihood of these separate forms of pretrial misconduct based on their one CPAT score. The new risk categories presented in the brief August report differentiate defendants into risk categories that are useful for pretrial decision-making.

With the exception of Age at First Arrest, all of the items are scored using a Yes-No scheme. Is this saying that for many of these items a defendant with multiple instances poses the same risk as defendants with just one instance (of a prior jail sentence, for example)?

Please see footnote 20 on page 15. Essentially, yes. It is possible that defendants with multiple instances do have a higher likelihood of pretrial misconduct than do defendants with only one instance, but the data do not support this at this time. Too few defendants have multiple instances to allow for statistical tests to be run on these data. This is why the defendant's score on the CPAT is just one of several items the court will consider when deciding what bond conditions may most likely mitigate an individual defendant's pretrial risk.

Has the CPAT been validated?

That depends on whom you ask. Many experts who develop risk assessment instruments in the pretrial field may say yes and many who do similar work in the corrections/recidivism field may say no. This difference results from how different experts define the term "validated." Despite this debate, the CPAT is an empirically-based pretrial risk assessment instrument that, because of the research and statistical methods used to develop it, enables all Colorado jurisdictions to use one risk assessment tool and thus better communicate about defendants' pretrial risk and coordinate services for defendants who have cases in multiple jurisdictions concurrently or over time. Furthermore, as jurisdictions implement the CPAT, they have been encouraged to collect data in a standardized and comprehensive manner so that the CPAT can be validated and/or re-validated in the future, as is the standard of practice for all risk assessment instruments.

Why was a 0.30 level of statistical significance used?

The 0.30 threshold is used only to include or exclude variables from the stepwise regression. This does not mean that all of the selected predictors have a significance level (p value) of 0.30 or less. The statistical output suggests that of the 12 predictors included in the final tool, most

have a very significant relationship with pretrial misconduct (i.e., higher than the standard 95% confidence level). Only a few predictors have a confidence level of 75% or 80%.

Some of the top 8 criminogenic risk factors that the LSI measures are not in the CPAT. Why are they risk factors on the LSI but not for pretrial misconduct?

The LSI was designed to measure risk and needs for criminal recidivism post-sentence. The CPAT was designed to measure both criminal activity (through the proxy of new filings) and failure to appear. So, they are measuring different things, on different populations (pre-sentence versus post-sentence), and for different time periods (the briefer pretrial period versus the longer post-sentence period). Also, different variables and/or different data definitions were used to develop the two tools, so the resulting tools will have different items.

It looks like data were collected on the effect of various bond conditions and different levels of pretrial services supervision. What did these data tell you?

The analyses to try to determine the extent to which the nature, level, dosage, or types of bond conditions and/or supervision affected pretrial success will be performed in phase 2 of the CISPR project, to the extent that the data allow for such analyses. The development of the risk tool was phase 1.

Why are the categories uneven in the quartile classification scheme (0-20, 21-30, 31-40, 41-100)? I get the 4 group idea, I just don't get why the groups are split the way they are (pg. 15).

The point range for the four groups is not the same because when all of the defendants were split into four nearly equally sized groups (i.e., quartiles), the range of points sometimes was more condensed and sometimes less condensed. That is, about 25% of defendants scored between 0-20, 25% scored between 21-30, 25% scored between 31-40, and 25% scored between 41-82. However, the four new risk categories described in the August 2012 brief report should be used.

When asking about prior arrests what is meant by arrest? Does it mean physically arrested or does it include summonses? Does arrest include traffic citations? If the source is the client, how does that person know what is meant? Why is there no verification, when such information is easily verifiable?

Please see the definition in Item 6 “Age at First Arrest” on page 16. If the defendant does not know what is meant by arrest, then the interviewer can clarify that arrest means when the defendant was first processed at a juvenile facility, taken into custody, or fingerprinted. This is how the question was posed to defendants during data collection and the item made the cut, so however defendants are answering this question is sufficient enough for the item to be included on the tool. As for verifying, juvenile arrest history is not always reliable in the online databases, and often does not include municipal arrests in which a juvenile is taken to the local police station or juvenile assessment center but not the county jail. In addition, self-reported age at first arrest is generally reliable as a criminal justice risk factor.²

² Bogue, B., Woodward, W., & Joplin, L. (2006). *Using a Proxy Score to Pre-screen Offenders for Risk to Reoffend*.

Is the public safety rate based on new arrests in Colorado or nationwide?

The public safety rate is based on new court filings instead of new arrests. The court filings pertain to criminal events that were alleged to have occurred in Colorado during the pretrial period of the defendant's case.

If alcohol use is a predictor, then why isn't drug usage?

Data on drug use or abuse was collected and analyzed multiple ways, and none of the ways produced findings that indicated drug use or abuse was statistically significantly related to pretrial misconduct in this study when the influence of other factors are accounted for. However, an alcohol and mental health item did make the cut.

After looking at the Appendix tables, I have a difficult time seeing how the factors actually chosen were chosen. What criteria were used? Was there a statistical cut off? If so, what was it? It seems that most of the factors on those page cluster fairly closely together. That's why I don't fully understand why some were chosen and some rejected.

Please see second paragraph on page 10 ("Data on many different variables...") for methods used to select what information was collected for the study. Please see the three paragraphs in the Bivariate Analysis section beginning on page 10. Corresponding footnotes 15 and 16 provide more information. Several factors are similar in content because the study's design intentionally included slight variations in several items (e.g., drug use/abuse) to maximize the chances that any factor might make the final cut.

Who else uses an empirically-developed pretrial risk assessment instrument?

There are many other jurisdictions that use empirically-developed pretrial risk assessment instruments, and the numbers are increasing each year. Like the CPAT in Colorado, some instruments were developed for use in multiple jurisdictions (e.g., Virginia, Ohio, Kentucky, Florida, Connecticut, Maine, the Federal court system), whereas others were developed for use in single jurisdictions (e.g., Hennepin County, MN; New York City, NY; Summit County, OH; Maricopa County, AZ).

Who thinks doing empirically-developed pretrial risk assessment is a good idea?

Both national criminal justice organizations and the general public support the use of empirically-developed pretrial risk assessment. The organizations that have recently issued a policy statement or resolution supporting the use of empirically-developed pretrial risk assessment include:

- National Association of Pretrial Services Agencies (national pretrial standards; 2004)
- American Bar Association (national pretrial standards; 2007)
- National Association of Counties (2010)
- American Probation and Parole Association (2010)
- American Jail Association (2011)

- Association of Prosecuting Attorneys (2011)
- American Council of Chief Defenders (2011)
- International Association of Chiefs of Police (2011)
- National Association of Criminal Defense Lawyers (2012)
- National Sheriffs' Association (2012)
- Conference of Chief Justices (forthcoming)

The general public also supports empirically-developed pretrial risk assessment. In 2012, the national polling firm Lake Research Partners found that over $\frac{3}{4}$ of called registered voters and focus group participants support using risk-based screening tools (“they said they are effective”) instead of cash bail bonds (“they said they are unjust”). Findings hold over differences in age, sex, conservative or liberal affiliation, and geographic region. Additionally for pretrial matters, voters said they would pay the most attention to judges and law enforcement officials, followed by crime victims’ groups and the sheriffs’ association. Bail bondsmen and the bail bond industry were last.