RISK ASSESSMENT OVERRIDES

Shuffling the Risk Deck Without Any Improvements in Prediction

THOMAS H. COHEN
CHRISTOPHER T. LOWENKAMP
Administrative Office of the U.S. Courts
KRISTIN BECHTEL
Arnold Ventures
ANTHONY W. FLORES
California State University, Bakersfield

In the federal supervision system, officers have discretion to depart from the risk designations provided by the Post Conviction Risk Assessment (PCRA) instrument. This component of the risk classification process is referred to as the supervision override. While the rationale for allowing overrides is that actuarial scores cannot always capture an individual's unique characteristics, there is relatively limited literature on the actual effects of overrides on an actuarial tool's predictive efficacies. This study examines overrides in the federal system by assessing the extent to which risk levels are adjusted through overrides as well as the impact of overrides on the PCRA's risk prediction effectiveness. Findings show that nearly all overrides lead to an upward risk reclassification, that overrides tend to place substantial numbers of persons under federal supervision (especially those convicted of sex offenses) into the highest supervision categories, and that overrides result in a deterioration of the PCRA's risk prediction capacities.

Keywords: supervision overrides; risk prediction; risk assessment tools; professional discretion

INTRODUCTION

Actuarial risk instruments have become a crucial element of the evidence-based community corrections paradigm (Andrews & Bonta, 2010a, 2010b; Andrews et al., 1990). Research has consistently shown that these tools provide a superior method of assessing the risk of recidivism for persons under supervision (hereinafter referred to as supervisees)

AUTHORS' NOTE: *Please note that a longer version of this article is available on the ResearchGate website. This article benefited from the careful editing of Ellen Fielding and Suzelle Fiedler. Correspondence concerning this article should be addressed to Thomas H. Cohen, Probation and Pretrial Services Office, Administrative Office of the U.S. Courts, One Columbus Circle, NE, Washington, DC 20544; e-mail: thomas_cohen@ao.uscourts.gov.*

CRIMINAL JUSTICE AND BEHAVIOR, 201X, Vol. XX, No. X, Month 2020, 1–21.
DOI: 10.1177/0093854820953449
Article reuse guidelines: sagepub.com/journals-permissions

compared with unstructured clinical approaches (Ægisdóttir et al., 2006; Grove et al., 2000; Schmidt et al., 2016). Moreover, actuarial tools are especially powerful because they provide community corrections professionals with a mechanism for making decisions about a supervisee's risk of recidivism that is both uniform and based on standardized sets of criteria. In the absence of these standardized criteria, professionals are often left with their subjective beliefs, notions, and biases to assess a supervisee's likelihood of committing new crimes, with the ultimate result being that professional judgment produces poorer risk predictions compared with actuarial instruments (Holsinger et al., 2001; McCafferty, 2015).

Although actuarial risk instruments outperform clinical prediction, most of these tools provide officers with an escape hatch that essentially allows them to disregard the instrument's designated risk classification and assign their own supervision level (McCafferty, 2015). For example, a supervisee classified as low risk by an actuarial instrument could be overridden to a higher supervision level should a corrections professional decide that, in his or her judgment, the actuarial risk designation underrepresents a supervisee's risk to reoffend (Cohen, Pendergast, & VanBenschoten, 2016). This aspect of the risk classification process is typically referred to as professional discretion or the supervision override and is one of the key components of the risk classification process (Andrews et al., 1990; McCafferty, 2015).

Those who favor using overrides assert that they are a necessary component when assigning supervision levels, because actuarial tools cannot always capture the unique characteristics of individuals that can be gleaned through an extensive investigation process (Cohen, Pendergast, & VanBenschoten, 2016). Hence, supervision overrides allow officers to depart from the actuarial risk classification when the totality of a supervisee's characteristics suggests that the person's adjusted supervision levels should diverge from the initial risk classification. By allowing officers to take into account the totality of a supervisee's characteristics, proponents of overrides assert that these mechanisms could potentially increase an actuarial tool's predictive accuracy (McCafferty, 2015). A growing but still relatively limited body of research, however, suggests that overrides could have the opposite result, diminishing a risk tool's predictive efficacy (Cohen, Pendergast, & VanBenschoten, 2016; McCafferty, 2015; Schmidt et al., 2016; Wormith et al., 2012).

This study attempts to augment the empirical literature on supervision overrides by analyzing this issue for the U.S. federal supervision system. Specifically, this study examines the use of supervision overrides for 259,571 persons on federal post-conviction supervision who were assessed with the Post Conviction Risk Assessment (PCRA). The PCRA is an actuarial instrument used by federal probation officers to assess the risk of recidivism for supervisees placed on federal terms of supervised release (i.e., TSR) or probation. By examining such a large population of supervised persons with risk assessments, this research can further enhance our understanding of overrides and their effect on risk prediction.

LITERATURE REVIEW

Since the establishment of the risk-needs-responsivity model, risk assessments have become a cornerstone for practitioners to target the risk and needs of their clients in an effort to improve the likelihood of success (Andrews & Bonta, 2010a, 2010b; Andrews et al., 1990). These tools are now infused into supervision practices expanding from juvenile through adult populations, and the process for implementing policies to inform decision-making based on the risk assessment output presents a number of valid questions about whether or not these policies (in this case, overrides) impact their validity.

While overrides are not directly synonymous with using exclusively professional or clinical judgment (presumably, practitioners are still following the risk assessment information to inform decision-making), it is still important to take stock of over 60 years' worth of research comparing clinical judgment to actuarial risk assessment. Overall, there is substantial and robust evidence to suggest that the use of unstructured clinical or professional judgment consistently results in weaker predictive validity in comparison to actuarial tools (see Ægisdóttir et al., 2006; Grove et al., 2000; Grove & Meehl, 1996; Harris, 2006). Despite this evidence, the challenge of adhering to or even trusting risk assessments persists. Meehl (1957), along with Dawes et al. (1989), described in great detail a number of the sources of resistance that commonly perpetuate the determination to ignore the actuarial assessment (Dawes et al., 1989, p. 11). These suspicious and perhaps critical viewpoints of risk assessment certainly persist among community corrections professionals (Clem, 2003; Schneider et al., 1996).

Some may suggest that this level of evidence should prompt policymakers to adopt solely actuarial processes (see Borum et al., 1993; Grove & Meehl, 1996; Quinsey et al., 1998, for more discussion). However, others argue that perhaps there is a need to strike a balance and determine if there is empirical value to integrating clinical judgment alongside actuarial decision-making (Gottfredson & Moriarty, 2006). One way community corrections professionals have attempted to weave discretion into actuarial decision-making is to allow officers authority to override or underride the results of the risk assessment when determining actual supervision levels.

While overrides have become a central facet of many risk assessment regimes, there is little empirical basis to support them (Cohen, Pendergast, & VanBenschoten, 2016; Wormith et al., 2012). Similar to the research consistently demonstrating that statistical prediction outperforms clinical judgment, the information on overrides, which could be characterized as a form of professional judgment, is beginning to reveal a similar trend—namely, the extensive use of overrides negatively impacts the predictive performance of risk assessments.

One common use of an override on supervision is based on charge type. Specifically, supervision clients who have convictions for sex crimes may receive an override that increases the intensity of supervision conditions. One such study aimed to test professional overrides and compare the predictive validity of the Level of Service/Case Management Inventory (LS/CMI) on a Canadian supervision sample comprising those convicted of sex crimes (N = 1,905) and convicted for general crimes (N = 24,245). The in-depth analysis of the initial and final risk levels by group revealed that officers were found to typically use overrides more frequently for clients convicted of sex crimes rather than general crimes (35.1% and 15.1% respectfully) and were more likely to increase, rather than decrease, risk level (14.9% and 1.6% respectfully). This override practice resulted in reductions in the predictive validity of the LS/CMI across all measures of recidivism (Wormith et al., 2012).

Schmidt et al. (2016) conducted a similar study examining professional overrides and comparing the predictive validity of the Youth Level of Services Case Management Inventory (YLS/CMI) on a sample of Canadian youth adjudicated with sex offenses (N = 204) and non-sex offenses (N = 184). When examining the frequency of professional overrides, the authors found that 41.6% of youth with non-sexual offenses in comparison to

74% of youth with sexual offenses experienced professional overrides, and all overrides increased the final risk level. In both samples, professional overrides decreased the predictive validity of the YLS/CMI risk levels for violent, non-violent, and sexual recidivism (Schmidt et al., 2016).

Several other studies on the LS/CMI and YLS/CMI that examine the influence of professional overrides on predictive validity are worth noting (Guay & Parent, 2018; Wormith et al., 2015). Wormith et al. (2015) sought to test the predictive validity of the LS/CMI and the impact of clinical overrides on a Canadian sample of Aboriginal (N = 1,692) and non-Aboriginal individuals (N = 24,758) that comprised those either on probation or released from a period of incarceration. Upon examination of the professional override, the authors found that predictive validity significantly decreased for general and violent recidivism in both the Aboriginal and non-Aboriginal samples when comparing the point-serial correlations for initial and final adjusted risk levels (Wormith et al., 2015).

In a more recent evaluation of the LS/CMI, Guay and Parent (2018) evaluated the impact of clinical overrides of the tool to adjust risk levels upward and downward on a Canadian sample (N = 3,646) of probation clients and those released from custody. Just 6.5% of the cases (N = 237) had a clinical override, with 144 cases resulting in a lower risk level and 93 cases with an increased risk level. For the sample that received the override, areas under the receiver operating characteristic (ROC) curve (AUC) values decreased between the original and final risk levels.

McCafferty (2015) tested the predictive validity and the influence of professional overrides on another instrument, the Ohio Youth Assessment System—Disposition tool (OYAS-DIS), on three measures of recidivism. Using a final weighted sample of 11,008 youth, the study results revealed that the OYAS-DIS original and adjusted risk levels were predictive of each recidivism measure. Overrides were observed in just 7% of the cases, but variation of the override rate was noted across Ohio counties. While the predictive validity decreased between original and adjusted risk levels, the results were not statistically significant.

Last, in a 2016 study examining the impact of overrides on the predictive validity of the PCRA, Cohen, Pendergast, and VanBenschoten pulled a sample of 58,524 initial PCRA assessments completed between August 2012 and December 2013. The findings indicated that 9.4% of the sample experienced an override, with low-risk cases having more overrides than high risk (Cohen, Pendergast, & VanBenschoten, 2016). However, when the authors examined recidivism rates comparing the initial and overridden risk levels, the results indicated that the cases with overrides recidivated at similar rates to their counterparts in the original risk level group (Cohen, Pendergast, & VanBenschoten, 2016).

Taken altogether, the research on overrides, while still relatively new, has observed that within both juvenile and adult settings and with some consistency, overrides tend to increase risk levels and supervision intensity, especially for specific offense types. When examining the predictive validity of the instruments evaluated, the research has indicated that overrides generally result in a decrease in the ability of the instrument to predict a variety of recidivism measures.

THE PRESENT STUDY

The present study attempts to augment the override research by examining several key issues including the overall prevalence of overrides for persons under federal supervision, the types of overrides (i.e., sex offender policy, other policy, or discretionary) utilized by officers, the rationales provided by officers when using discretionary overrides, and the adjustments in risk levels that occur as a result of overrides. Importantly, we also explore whether overrides are associated with a degradation in the PCRA's risk prediction capacities. The research questions are detailed below.

Research Question 1: How frequently are overrides used in the federal supervision system? Which types of overrides are employed most frequently? What are the rationales officers provide when using discretionary overrides?

Research Question 2: To what extent are the original PCRA risk classifications adjusted by supervision overrides? Are certain types of overrides (e.g., sex offender policy overrides) more likely than other types of overrides (e.g., discretionary overrides) to move supervisees across multiple supervision levels?

Research Question 3: What is the association between supervision overrides and the PCRA's capacity to predict recidivism among supervisees? Compared with the PCRA's original risk levels, how different are the adjusted risk levels in terms of predicting recidivism outcomes? How does the PCRA's predictive performance differ among comparable groups of supervisees with and without overrides?

METHOD

PARTICIPANTS

Data for this study were obtained from 94 U.S. federal judicial districts and comprised 259,571 supervisees with initial PCRA assessments conducted between August 31, 2012, and July 30, 2017. We ended the data extraction date in mid-2017 because the PCRA's supervision levels were modified to include a violence trailer at that time (Serin et al., 2016). Regardless, this study sample constitutes one of the largest ever used to examine the presence of overrides among supervisees and the association between overrides and risk prediction (Chappell et al., 2012; Guay & Parent, 2018; McCafferty, 2015; Vaswani & Merone, 2014; Wormith et al., 2012). Table 1 provides a descriptive overview of the overall study population. About a third of the study population comprised either non-Hispanic Whites (35%) or Blacks (35%), while Hispanics of any race accounted for about a quarter (24%) of the population. Females comprised 18% of the study population, and the average supervisee age was about 40 years (*SD* = 11.7). Regarding the supervisee's PCRA risk classifications, 32% were classified into the low-risk category, 38% were deemed low/moderate risk, 22% were designated moderate risk, and 9% were placed in the highest risk category.

MEASURES

Overrides and the PCRA

Judicial policy provides federal probation officers with discretion to depart from the PCRA's initial risk designations and place supervisees into higher or lower supervision categories (Guide to Judiciary Policy, 2014). The PCRA is a fourth-generation risk tool and is used to assess a supervisees' risk of recidivism through a process in which federal probation officers score supervisees on 15 static and dynamic risk predictors. These 15 predictors are also used to generate a raw PCRA score ranging from 0 to 18, which translates into the following four risk categories: low (0–5 points), low/moderate (6–9 points), moderate

TABLE 1: Descriptive Statistics in Study Sample

			Override types		
Variables	Full sample	Any override	Policy sex offender	Policy other[a]	Discretionary
N (% of sample)	259,571	27,480 (10.6)	12,321 (4.8)	1,931 (0.7)	13,228 (5.1)
Average age (*SD*)	39.8 (11.7)	43.1 (12.5)	44.7 (13.0)	41.9 (11.6)	41.7 (11.9)
Race (%)					
White	35.2	54.2	68.1	38.7	43.5
Black	34.5	24.7	14.2	38.1	32.5
Hispanic, any race	24.1	14.1	8.8	14.9	18.9
Other	6.1	7.0	8.9	8.3	5.0
Female (%)	17.7	11.2	1.9	19.6	18.6
Offense type (%)					
Drugs	43.8	23.0	5.5	25.9	38.9
Financial offenses	20.3	13.5	2.6	15.5	23.4
Weapons/firearms	14.8	9.9	5.1	22.3	12.6
Violence	7.3	16.0	19.5	25.6	11.3
Immigration/customs	4.7	2.0	0.7	3.6	3.0
Sexual offense	3.1	26.3	54.1	2.1	4.0
Obstruction/escape	2.5	6.9	11.7	3.0	3.0
Traffic/DWI	2.2	1.2	0.0	0.4	2.4
Public order	1.2	1.2	0.9	1.7	1.5
PCRA risk classification (%)					
Low	31.7	46.5	45.4	16.5	51.8
Low/moderate	37.8	34.4	34.1	42.7	33.5
Moderate	22.0	19.0	20.5	40.8	14.5
High	8.5	0.1	0.0	0.0	0.2
Average PCRA score (*SD*)	7.4 (3.7)	6.3 (3.2)	6.2 (3.3)	8.5 (2.8)	6.1 (3.1)
Arrested within 24 months of initial assessment (%)	18.9	15.0	13.6	23.8	15.0

Note. Includes supervisees on federal post-conviction supervision or probation with initial PCRA assessments whose arrest activity could be tracked for a minimum of 2 years. PCRA = Post Conviction Risk Assessment.
[a]Includes policy overrides that occurred because the supervisee was either mentally ill, violent, or is a serious youthful supervisee.

(10–12 points), or high (13 or more points). Unless otherwise noted, these risk categories have been dummy coded into the following values in the current study (0 = low supervision, 1 = low/moderate supervision, 1 = moderate supervision, and 1 = high supervision). Information about the PCRA's history, development, risk scoring scales, and predictive validity can be obtained through a variety of published articles and technical reports (see Cohen, Lowenkamp, & VanBenschoten, 2016; Cohen & VanBenschoten, 2014; Lowenkamp et al., 2013, 2015; Luallen et al., 2016; Skeem & Lowenkamp, 2016).

When using the PCRA to assess a supervisee's risk level, judicial policy provides officers with discretion to override these persons into alternative supervision levels if they think, in their own professional judgment, that the PCRA's risk score under- or overrepresents a supervisee's risk to reoffend (Cohen, Pendergast, & VanBenschoten, 2016). Judicial policy also states that overrides should be relatively rare and that officers should use overrides for only certain case types (i.e., policy overrides) or supply rationales for employing overrides (i.e., discretionary overrides). Policy overrides involve instances where officers move supervisees into higher or lower supervision levels because they meet the following criteria: (a) are classified as persons convicted of sex offenses, (b) evidence patterns of persistently violent behavior, (c) manifest indications of severe mental illness, or (d) are youthful supervisees with extensive criminal histories. In addition to policy overrides, judicial policy provides officers latitude to issue overrides for other reasons, which are called discretionary overrides. A comprehensive justification is required whenever the officer decides to override a supervisee for discretionary reasons. Any override request must be reviewed and approved by a supervisory officer (Administrative Office of the U.S. Courts [AOUSC], 2011).

Outcome measures

The primary outcome of interest involves whether a supervisee was rearrested within 24 months of their initial PCRA assessment. Rearrests for new criminal activity, which we also refer to as recidivism, were obtained from the National Crime Information Center (NCIC) and Access to Law Enforcement System (ATLAS). ATLAS is a software program used by the AOUSC that provides an interface for performing criminal record checks through a systematic search of official state and federal rap sheets. Recidivistic events were defined to include arrests for either felony or misdemeanor offenses (excluding arrests for technical violations) within 24 months of the initial PCRA assessment.

ANALYTICAL PLAN

To examine the presence of overrides for supervisees on federal post-conviction supervision and the potential of overrides to degrade the PCRA's predictive capacities, we calculated descriptive statistics, effect sizes, and measures of predictive discrimination (e.g., areas under the receiver operating characteristic [ROC] curve [AUCs]). Specifically, we compared the AUC-ROC scores and used chi-square tests to assess whether the original as opposed to adjusted PCRA risk levels performed better in terms of predicting recidivism.

In addition to examining differences in prediction of pre- and post-supervision adjustment, we explored whether the PCRA's predictive capacities differed when comparing supervisees whose supervision levels remained unchanged with similarly situated supervisees whose risk levels were adjusted through an override. When comparing the PCRA's predictive efficacy for supervisees with and without overrides, it is important to acknowledge that supervisees with overrides differ in many ways from supervisees without overrides on key domains associated with recidivism (see Table 1 and the appendix). Accounting for these differences is crucial when attempting to assess whether adjustments for supervisees with overrides resulted in declines in the PCRA's risk prediction compared with similarly situated supervisees without overrides. We address this issue by employing propensity score matching (PSM) techniques to generate matched groups of supervisees with and without supervision overrides. PSM has become a commonly employed technique to estimate treatment effects when randomized assignment is unavailable and it becomes necessary to account for covariates that could influence the outcome of interest (Rosenbaum & Rubin, 1983, 1985).

In this study, the "treated" supervisees constituted those who received overrides. Supervisees were matched on their PCRA criminal history, education/employment, substance abuse, social networks, and cognition domain scores. They were also matched by their demographic characteristics (i.e., age, race/ethnicity, and sex) and by the federal judicial district where they were supervised. With exception of the PCRA domains, which are numeric variables with values ranging from 0 to 9 depending upon the specific domain examined, the demographic and district variables were captured using dummy variable coding (0/1). We also explored matching by the instant conviction offense, but decided against doing so because of the large number of persons convicted of sex offenses who would have been lost through matching.

We performed three matches to account for the override types. The first created matched groups of supervisees who did and did not receive any override; second, we matched by whether a supervisee received a policy sex offender override; and third, we matched by whether a supervisee received a discretionary override. We did not perform matching for the "other" policy overrides because of the relatively few supervisees in this group, though they were included in the any override category. A two-step process was employed for all three of the matching models. First, we estimated propensity scores using logistic regression in which we predicted the likelihood of a supervisee receiving an override. Then, we used the estimated likelihood scores to match the override group to a comparison group of supervisees without overrides, applying one-to-one nearest neighbor matching with a .0001 caliper setting (Guo & Fraser, 2014). This method resulted in matches being found for 92% of supervisees with any overrides, 85% of supervisees with sex offender overrides, and 94% of supervisees with discretionary overrides.

An example of the results from the matching procedures is shown in the appendix displaying the balance obtained between the any override and non-override supervisees. This appendix shows that the matching procedures generated override and non-override groups of supervisees, showing strong balance on the key covariates of interest. After matching, we employed multivariate logistic regression to test for slope and intercept differences, measuring the association between the supervision categories and recidivism outcomes for supervisees with and without overrides. Last, we used a more conservative alpha level of .001 to denote statistical significance and report effect sizes whenever possible because of the large sample sizes analyzed in this study.

RESULTS

SAMPLE DESCRIPTION

The first goal in this research was to understand the ways in which cases given an override differ from those where the original classification and assigned supervision level match. Table 1 contains the descriptive statistics on supervisee and offense characteristics, and risk information and outcomes for the total sample, as well as the sample of cases that were assigned an override. Overall, 10.6% of the sample was given an override for one of the various justifications for an override. More specifically, 4.8% of the sample received an override because of being identified as a person convicted of sex offenses, 0.7% of the sample received an override for other policy reasons, and 5.1% of the sample received an override for discretionary reasons. The average age of the total sample is somewhat lower than the age for those supervisees given an override. Moreover, White supervisees make up larger proportions of overrides for sex offenses and discretionary overrides relative to the proportion of White supervisees in the overall sample, while Blacks seem to be represented fairly equally in "other" policy overrides and discretionary overrides but are far less likely to be given an override for sex offending. Regarding the risk characteristics and recidivism

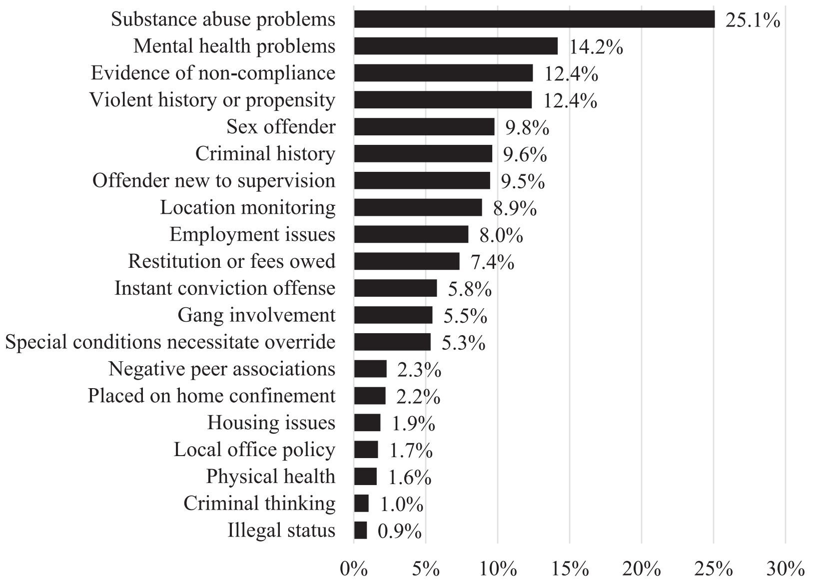

Figure 1: Rationales Offered by Officers for Discretionary Overrides

Note. The rationales provided for discretionary overrides were identified and coded from 11,502 of the 13,228 discretionary overrides. Totals do not sum to 100% as officers can provide multiple rationales for overriding supervisees.

rates, 46.5% of supervisees given any override were categorized as low risk. Smaller percentages of low/moderate, moderate, and high-risk categories are present in the override sample when compared with the total sample. Finally, the override sample has lower recidivism rates than the entire sample.

Of interest is the reason behind discretionary overrides, as officers are able to use a number of different reasons for applying a discretionary override (see Figure 1). As is indicated, substance abuse problems are the most common justification for a discretionary override (25.1%). Mental health problems, evidence of noncompliance, and violent history or propensity make up fairly equal percentages of justifications (14.2%, 12.4%, 12.4%, respectively). Having a history of sex offending (9.8%), criminal history (9.6%), being new to supervision (9.5%), and having a location monitoring condition (8.9%) also make up similar percentages of the discretionary overrides. It is worth noting that some of the most common justifications for discretionary overrides are covered by the items on the PCRA.

ORIGINAL AND OVERRIDE RISK CATEGORIES

Data distinct to the sample that received an override are contained in Table 2. These data present the number of risk levels changed via override for all overrides as well as by each type of override. In Table 2, the data for all overrides are presented. Of the cases identified as low risk on the PCRA, 42.7% were reclassified as low/moderate risk, 10.1% were reclassified as moderate risk, and 47.3% were reclassified as high risk. Cases identified as low/moderate risk based on the PCRA score were adjusted to moderate risk 37.0% of the time and to high risk 61.6% of the time. A small percentage (1.4%) were overridden to the low-risk category.

TABLE 2: Change in Risk Levels for Supervisees With Overrides, by Override Type and Initial Risk Level

PCRA risk categories before override and override types	Number supervisees	Override change score (%)			
		− 1 or more risk levels	+ 1 risk level	+ 2 risk levels	+ 3 risk levels
Any override					
Low	12,766	0.0	42.7	10.1	47.3
Low/moderate	9,452	1.4	37.0	61.6	0.0
Moderate	5,229	1.7	98.3	0.0	0.0
High	33	100.0	0.0	0.0	0.0
Policy-sex offender					
Low	5,596	0.0	3.5	6.8	89.7
Low/moderate	4,201	0.0	5.7	94.3	0.0
Moderate	2,523	0.0	100.0	0.0	0.0
High	1	—	0.0	0.0	0.0
Policy-other					
Low	319	0.0	43.0	27.9	29.2
Low/moderate	825	0.0	43.4	56.6	0.0
Moderate	787	0.1	99.9	0.0	0.0
High	0	—	—	—	—
Discretionary					
Low	6,851	0.0	74.6	11.9	13.5
Low/moderate	4,426	3.1	65.4	31.5	0.0
Moderate	1,919	4.4	95.6	0.0	0.0
High	32	100.0	0.0	0.0	0.0

Note. Includes supervisees who received supervision overrides. "—" = Not applicable or too few supervisees to produce statistically reliable estimates; PCRA = Post Conviction Risk Assessment.

When turning to sex offender policy overrides, a different trend is noted. Almost all of the sex offender policy overrides move the supervisees to the high-risk category. Of the low-risk supervisees, 89.7% are reclassified as high risk. A larger percentage (94.3%) of the low/moderate risk supervisees are reclassified as high risk and all the moderate-risk supervisees are reclassified as high risk. Policy overrides for reasons other than convicted sex offense status typically involve an increase of one or two risk levels. However, with low-risk supervisees, where an increase in three risk levels is possible, 29.2% of the overrides reach this threshold. Discretionary overrides show a somewhat different trend. Low-risk supervisees with discretionary overrides tend to be moved up one risk level (74.6% are overridden to low/moderate risk). This is true with low/moderate risk supervisees and moderate risk as well (65.4% and 95.6% respectively). In the latter case, however, there is only one subsequent risk level, and it is unknown if that group of supervisees would be overridden more than one category if a risk category above "high" existed.

RECIDIVISM RATES BY ORIGINAL AND OVERRIDE RISK CATEGORIES

The recidivism rates for any offense within 24 months of initial assessment for supervisees with overrides by risk category are contained in Table 3. The results are presented for original PCRA risk classifications as well as the adjusted risk levels. The data are presented in four panels and display recidivism rates for all overrides, those supervisees with adjusted risk levels due to sex offender policy overrides, other policy overrides, and discretionary overrides.

TABLE 3: Recidivism Rates Any Offenses (24 Months of Assessment) by Original and Adjusted Risk Levels

	Original PCRA risk levels		Adjusted PCRA risk levels		
Risk levels & override types	Number	Arrested (%)	Number	Arrested (%)	Chi-square
Any override					
Low	12,766	7.0	175	9.7	729.4*
Low/moderate	9,452	18.7	5,503	7.5	
Moderate	5,229	27.9	4,801	16.9	
High	33	18.2	17,001	16.9	
Total	27,480	15.0	27,480	15.0	
AUC-ROC, 99.9% CI	0.67 [0.66, 0.68]		0.56 [0.55, 0.57]		
Policy-sex offender					
Low	5,596	6.5	0	—	528.0*
Low/moderate	4,201	16.2	196	6.6	
Moderate	2,523	25.0	623	10.8	
High	1	—	11,502	13.8	
Total	12,321	13.6	12,321	13.6	
AUC-ROC, 99.9% CI	0.66 [0.64, 0.69]		0.51 [0.50, 0.52]		
Policy-other					
Low	319	8.2	0	—	11.4*
Low/moderate	825	20.6	138	5.1	
Moderate	787	33.6	447	15.4	
High	0	—	1,346	28.5	
Total	1,931	23.8	1,931	23.8	
AUC-ROC, 99.9% CI	0.63 [0.59, 0.68]		0.60 [0.56, 0.63]		
Discretionary					
Low	6,851	7.4	175	9.7	96.1*
Low/moderate	4,426	20.6	5,169	7.6	
Moderate	1,919	29.3	3,731	18.1	
High	32	18.8	4,153	21.7	
Total	13,228	15.0	13,228	15.0	
AUC-ROC, 99.9% CI	0.67 [0.65, 0.69]		0.63 [0.61, 0.65]		

Note. Table includes supervisees with policy or discretionary overrides. "—" = Not enough cases to generate statistically reliable estimates; PCRA = Post Conviction Risk Assessment; AUC-ROC = area under the receiver operating characteristic curve; CI = confidence interval.
*$p < .001$.

Across each of the categories in Table 3, an interesting trend is noted. Specifically, it is apparent that the failure rates of the original PCRA risk categories increase monotonically from low through moderate risk. The exception is high risk; however, very few supervisees with overrides are initially categorized as high risk ($n = 33$ supervisees). Furthermore, the overall failure rates for all overrides, sex offender policy overrides, and discretionary overrides are fairly similar at 15.0%, 13.6%, and 15.0%. The other policy overrides, however, have a somewhat higher overall failure rate of 23.8%. Finally, when reviewing the data on the original PCRA classifications, note that the AUC-ROC values range from 0.63 (other policy overrides) to 0.67 (any override and discretionary overrides).

The recidivism rates for any offense based on the adjusted PCRA levels indicate that while the recidivism rates increase as risk category increases, these increases in failure rates are somewhat "flatter" than those observed with the original PCRA levels. In addition, it is the case with discretionary and all overrides that the recidivism rates for low/moderate risk supervisees are lower than those for low-risk supervisees. Also of note is the fact that the AUC-ROC values decrease compared with those generated from the original PCRA categories and range in value from 0.51 to 0.63.

While it might be tempting to conclude that the lower recidivism rates for supervisees reclassified to higher risk levels is the result of effective supervision, note that the overall recidivism rates do not differ (e.g., all overrides recidivism rate equals 15.0% for the original PCRA classification and the adjusted classification, even though the recidivism rate for moderate risk is 27.9% and 16.9% for the original and adjusted PCRA categories respectively). In other words, the overall failure rates remain constant, because adjusting risk levels merely leads to a "re-shuffling" of the same cases into different categories. These adjusted risk categories manifest impoverished risk prediction compared with the original risk groupings.

MULTIVARIATE LOGISTIC REGRESSION MODELS WITH MATCHED SAMPLE

Table 4, along with Figures 2 to 4, explore the relationship between supervision overrides and the PCRA's recidivism risk prediction efficacy for any offenses through a mixture of statistical controls, including propensity score matching and binary logistic regression (for further details, see "Method" section). The first regression models—no interaction terms—detailed in Table 4 show supervisees with overrides having significantly lower arrest likelihoods for any offenses compared with their matched counterparts without overrides. Specifically, the odds of recidivism for crimes involving any behavior were about 60% lower ($p < .001$) for supervisees with overrides than for similarly situated supervisees with unadjusted risk levels. This finding of lower intercepts regarding the outcome of interest for supervisees with overrides held irrespective of the type of override (i.e., any override, policy sex offender override, or discretionary override). Clearly, the decision to override a supervisee because of perceived notions of higher levels of recidivistic activity was not reflected in the supervisee's actual rearrest behavior, which was uniformly lower than that of supervisees without overrides.

The second set of models in Table 4 utilizes interaction terms to investigate the extent to which a supervisee's PCRA supervision levels and rearrest odds are moderated by overrides. Results show that the form of the relationship between the PCRA's supervision levels and a supervisee's rearrest activity for any offenses differs significantly between supervisees without overrides and supervisees receiving any overrides, $\Delta\chi^2(3) = 44.8$; $p < .001$, or discretionary overrides, $\Delta\chi^2(3) = 38.6$; $p < .001$. Conversely, supervisees with policy sex offender overrides did not manifest significantly divergent patterns of rearrest activity across the PCRA supervision levels, compared with similarly situated supervisees without supervision adjustments, $\Delta\chi^2(2) = 6.3$; *ns*. For the sex offender models, supervisees designated low risk were combined into the low/moderate risk category as there were basically no supervisees with a policy sex offender override who remained low risk after the override adjustment (see Table 4 and Figure 3).

While the tables demonstrate significant interaction effects, how the likelihood of rearrest is influenced by overrides across the four PCRA supervision levels is better demonstrated through a series of predicted probabilities highlighted in Figures 2 to 4. The predicted probabilities were calculated based on the logistic regressions with interactions for all the override types (i.e., any, policy sex offender, and discretionary).

TABLE 4: Logistic Regression Examining the Relationship Between PCRA Supervision Levels and Arrest Within 24 Months of Assessment (Any Offense) for Matched Supervisees With and Without Overrides

	Model 1			Model 2		
		99.9 CI			99.9 CI	
Variables & override types	Odds ratio	Lower	Upper	Odds ratio	Lower	Upper
Any override						
Any override	0.43*	0.38	0.50	1.90	0.74	4.86
PCRA supervision levels						
Low/moderate	3.41*	2.91	4.00	3.56*	3.01	4.21
Moderate	7.20*	6.02	8.60	6.71*	5.52	8.17
High	8.74*	7.05	10.84	10.66*	8.40	13.53
Override × PCRA Supervision						
Override × Low/Moderate	—	—	—	0.21*	0.08	0.54
Override × Moderate	—	—	—	0.28*	0.11	0.72
Override × High	—	—	—	0.19*	0.07	0.49
Constant	0.06			0.06		
Model chi-square	1504.7			1594.8		
Number of supervisees	50,360			50,360		
Policy sex offender						
Policy sex offender override	0.30*	0.23	0.41	0.59	0.24	1.44
PCRA supervision levels						
Moderate	3.64*	2.93	4.53	3.68*	2.93	4.61
High	4.76*	3.51	6.46	4.92*	3.57	6.76
Override × PCRA Supervision						
Override × Moderate	—	—	—	0.50	0.17	1.47
Override × High	—	—	—	0.50	0.19	1.29
Constant	0.12			0.12		
Model chi-square	534.5			557.8		
Number of supervisees	20,514			20,514		
Discretionary override						
Discretionary override	0.49*	0.41	0.58	1.50	0.61	3.70
PCRA supervision levels						
Low/moderate	2.68*	2.15	3.33	3.00*	2.40	3.74
Moderate	5.70*	4.50	7.22	5.23*	4.12	6.63
High	8.01*	6.15	10.44	8.51*	5.72	12.66
Override × PCRA supervision						
Override × Low/Moderate	—	—	—	0.26*	0.10	0.67
Override × Moderate	—	—	—	0.40*	0.16	0.96
Override × High	—	—	—	0.31*	0.10	0.93
Constant	0.07			0.07		
Model chi-square	1007.8			1054.7		
Number of supervisees	24,942			24,942		

Note. Standard errors for model (not shown) clustered at the district level. "—" = Not applicable; PCRA = Post Conviction Risk Assessment; CI = confidence interval.
$^*p < .001$.

According to Figure 2, the predicted probabilities of arrests take different forms among the four PCRA supervision categories depending upon whether they involved overrides. For example, the predicted probability of arrest increases in a monotonic fashion for supervisees whose risk levels remained unadjusted at the time of initial assessment. In comparison, the

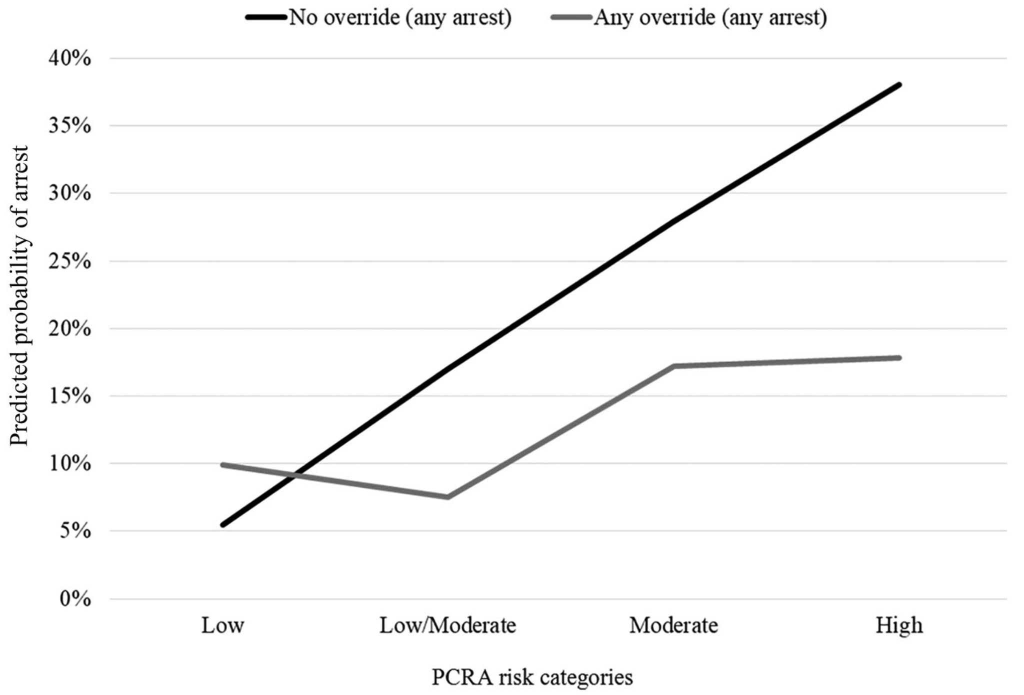

Figure 2: Predicted Probabilities of Arrest (Any Offense) Within 24 Months of Assessment for Matched Supervisees With and Without Any Supervision Overrides
Note. PCRA = Post Conviction Risk Assessment.

predicted probabilities of arrest for the matched group of supervisees with overrides differs substantially from those of their non-override counterparts. Specifically, the override group manifests a more leveled slope across the PCRA supervision categories. Initially, the predicted arrest probabilities decline from the low to the low/moderate supervision categories and then increases slightly between supervisees designated low/moderate to moderate risk. Afterward, the predicted arrest probabilities remain essentially unchanged for supervisees supervised at the moderate- or high-risk categories. A pattern similar to the any override figure manifests itself for supervisees with discretionary overrides (see Figure 4). For supervisees with sex offender overrides, the probability of arrest rises simultaneously with risk; however, the increase in slope is relatively gradual, and almost flat, compared with supervisees matched on similar characteristics without overrides (see Figure 3).

DISCUSSION

The current study examined professional overrides for supervisees under federal post-conviction supervision. About 11% of 259,571 supervisees on post-conviction supervision received an override, with slightly over half of these adjustments involving policy rather than discretionary overrides. Most policy overrides (86%) involved sex offender rather than "other" policy overrides. In terms of the effect of overrides on actual supervision levels, nearly all the overrides adjusted the risk levels upward, with supervisees being placed into higher risk levels than they would have been had the officers followed the original PCRA

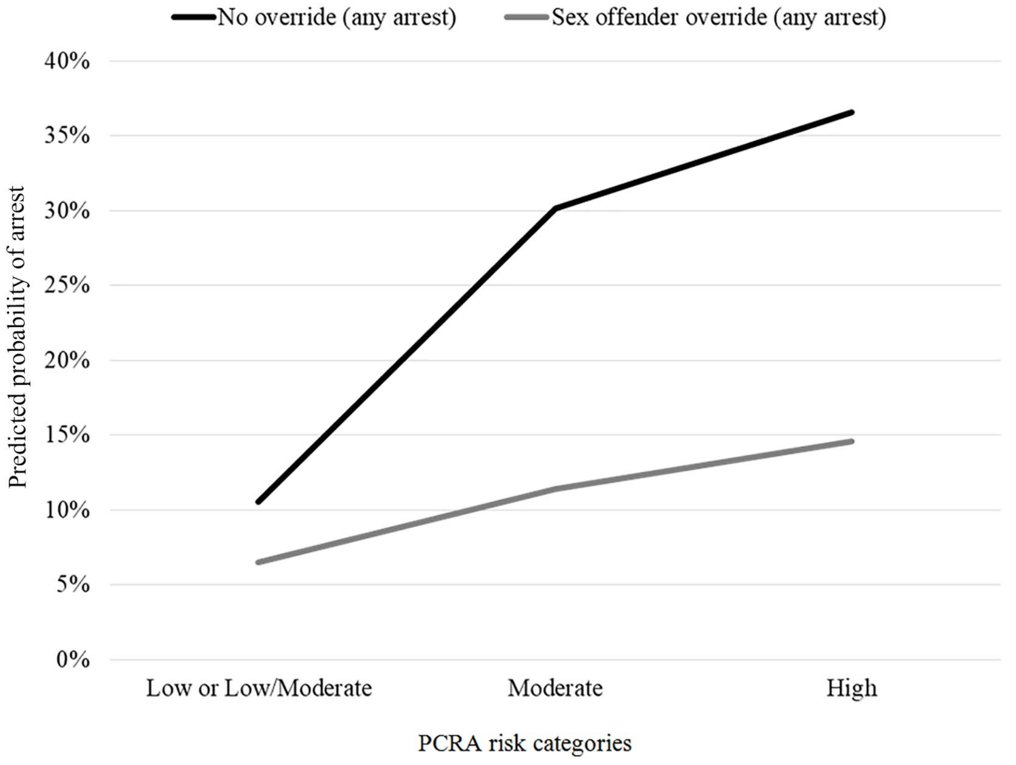

Figure 3: Predicted Probabilities of Arrest (Any Offenses) Within 24 Months of Assessment for Matched Supervisees With and Without Sex Offender Policy Overrides

Note. PCRA = Post Conviction Risk Assessment. The sex offender models combine supervisees on low or low/moderate supervision.

risk designations. Supervisees with policy-sex offender overrides received the most substantial adjustments; nearly all (93%) of these supervisees were reshuffled into the highest risk levels. The reshuffling of supervisees with other policy or discretionary overrides was less considerable; these supervisees were typically reclassified into a risk category one level higher than their original risk level.

In general, this reshuffling of supervisees from lower to higher risk categories did not improve the PCRA's risk prediction capacities; rather, it weakened them. This diminishment in risk prediction can be observed by the fact that the adjusted risk levels performed worse than the original risk levels in terms of predicting the likelihood of recidivism. The deterioration in risk prediction was especially striking for supervisees with upward adjustments resulting from sex offender overrides. By placing nearly all persons convicted of sex offenses into the highest risk category, the resultant outcome is that these persons are essentially treated the same in terms of their likelihood of recidivism. Consequently, the adjusted risk levels resulting from this type of override produce accuracy metrics (e.g., AUC scores) indicating relatively poor performance. Although the discretionary and other policy overrides tended to reshuffle supervisees less dramatically, these overrides also resulted in a degradation in the PCRA's risk prediction capacities.

The pattern of the adjusted risk levels performing poorer at prediction compared to the original risk levels classifications is probably best illustrated in Figures 2 to 4, which show

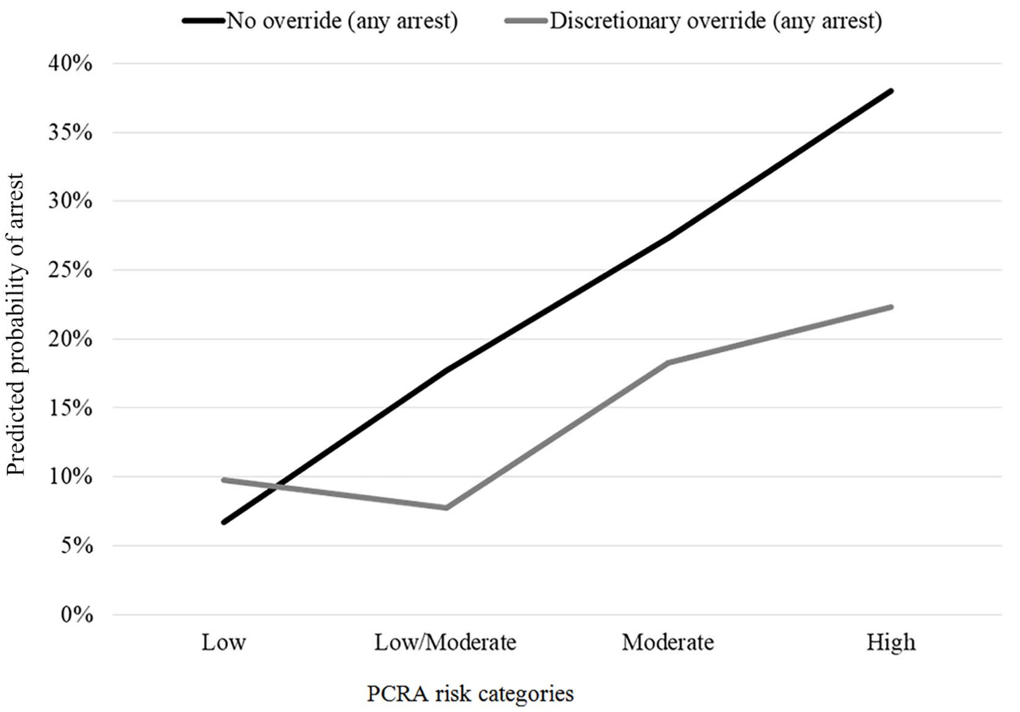

Figure 4: Predicted Probabilities of Arrest (Any Offenses) Within 24 Months of Assessment for Matched Supervisees With and Without Discretionary Overrides

Note. PCRA = Post Conviction Risk Assessment.

a "flattening" of the failure rates across the four risk levels compared with similarly situated supervisees whose risk levels remained unadjusted. Ultimately, the study's results support the contention that overrides lead to a reshuffling of supervisees into higher risk categories without any concomitant improvement in prediction.

Our findings that overrides result in an upward reshuffling of supervisees into higher risk levels, especially for persons convicted of sex offenses, while producing a simultaneous degradation in risk prediction are consistent with other studies focusing on supervision overrides (see Cohen, Pendergast, & VanBenschoten, 2016; Guay & Parent, 2018; McCafferty, 2015; Schmidt et al., 2016; Vaswani & Merone, 2014; Wormith et al., 2015). The study's principal findings, in alignment with other research, that supervision overrides result in the degradation of a risk tool's predictive capabilities is particularly important for community correction professionals. Specifically, why should systems that have adopted the risk, needs, and responsivity (RNR) model allow for professional overrides at all, and what set of circumstances leads officers to place supervisees into higher risk levels than those designated by their actuarial tools?

In terms of systemwide approval of the use of overrides, until recently, the federal supervision system allowed officers at their discretion to override persons convicted of sex offenses into the highest risk levels regardless of their initial risk classification. Allowing officers' discretion to apply overrides for supervisees convicted of or with a history of sex offenses was primarily based on rationales involving the capacity of generalized risk instruments like the PCRA to predict recidivistic outcomes for persons convicted of sex offenses as well as community safety concerns. The rationale that risk instruments may have poor predictive capacities for the convicted sex offense population was based on the unproven idea that generalized risk instruments might not be viable for this specific subpopulation supervisees. Moreover, many officers maintain that persons convicted of sex offenses are simply untreatable and, hence, that the system should "err on the side of caution" to protect community safety.

Whether federal supervisees convicted of sex offenses should be extensively subjected to overrides has been brought into question by recent research showing that the PCRA can, in fact, predict general and violent recidivism for these supervisees (Cohen, 2018) and that many of these supervisees are rearrested at rates below those of the general supervision population (Cohen & Spidell, 2016). As a result of these studies, federal supervision policies were recently changed so that officers now must hold in abeyance the assignment of actual supervision levels until a thorough background investigation has been completed. Since this change in supervision policy occurred after these data were extracted, future studies will have to assess whether officers have altered the way they use overrides for this group of supervisees.

Outside sex offender overrides, this research raises the issue of what leads officers to override supervisees into higher risk categories. The extant documentation of the rationales cited when using discretionary overrides provides some intriguing insights into this issue. Specifically, over half the rationales used for discretionary overrides involve factors already measured by the PCRA, including substance abuse problems, evidence of noncompliance, criminal history, employment issues, and negative peer associations. Moreover, an additional 36% of discretionary overrides encompassed issues that should place these overrides into the policy sex or policy "other" override category.

Reviewing these rationales provides a glimpse into the role that professional discretion might play when officers decide to move supervisees into higher risk categories. In cases where the officers cite risks mirroring those already assessed by the PCRA, it may be that officers believe that, although the tool purports to measure risk factors like substance abuse problems or criminal history, it does not sufficiently weight these factors. Instead, the officers seem by the overrides to be registering their opinion that the PCRA has not adequately measured a supervisee's substance abuse problems or criminal record and that, in their professional opinion, these issues should be given greater weight when assessing a supervisee's risk of recidivism and appropriate management levels. The documentation officers provided in the override text illustrates this view, as they highlight issues such as extensive abuse and/or addiction to various illicit substances and the failure of many of these supervisees to seek adequate treatment.

In terms of what explains officers' behavior when citing to rationales that involve policy overrides (e.g., sex offender, mental health, propensity toward violence), a close examination of these justifications suggests that officers may be implying that using checkboxes, even when applying them in the context of policy overrides, somehow neuters their autonomy or substitutes their ability to justify their decisions when moving supervisees into higher risk categories. Officers, for example, would write extensive summaries of a supervisee's mental health issues when citing to that particular rationale for an override. Again, writing such extensive summaries seems to imply that checking a "mental health" override box inhibits an officer's ability to adequately assess, detail, and highlight the mental health problems and challenges supervisees face when starting supervision.

Regardless of whether the override rationales involved a reiteration of factors already measured by the PCRA or issues covered by policy overrides, either justification essentially amounts to a "double-counting" of risk factors in the override decision (Chouldechova, 2020; Hamilton, 2015). Such forms of double-counting should be discouraged because they amount to the same factors being considered multiple times when deciding whether to place a supervisee into higher supervision levels with greater levels of supervision intensity. These types of decisions multiply a single incident "without the decision maker necessarily being cognizant of the overlap. The risk prediction will likely be higher than appropriate and the consequences to the individual may also be magnified" (Hamilton, 2015, pp. 97–98).

Last, it is important to note that several of the rationales cited by officers, including location monitoring or the collection of fines/fees/restitution, are indicative of increased workload or case activity. Essentially, these overrides occurred not because officers, in their opinion, decided that a supervisee's risk is higher than that assessed by the PCRA, but because certain conditions (e.g., location monitoring) imposed on supervisees necessitated more intensive workload or supervision requirements. Whether overrides should occur because of perceived workload demands is questionable and should be a topic of discussion within the federal supervision system.

Although this study could be used to support the contention that the federal supervision system might be better off disallowing overrides, prohibiting their use could potentially have negative implications for the federal supervision system. Officers may balk at adhering to the PCRA's risk classifications if they conclude that their professional judgment has been completely superseded by the actuarial tool (McCafferty, 2015). Officers might, for instance, oversee supervisees placed into the lowest risk categories at higher levels of intensity than should be warranted. Officers, moreover, could subvert the risk classification process by manipulating the scores to place supervisees into higher risk categories than would occur if accurate scoring techniques were applied (Gebo et al., 2006; McCafferty, 2015).

These scenarios suggest that overrides will always have a place in actuarial systems. Essentially, overrides provide a way for officers to continue applying their professional discretion so that assessments serve the purpose of informing rather than mandating how supervisees should be managed. However, there is a need to minimize their use. The current and prior research showing diminishment in prediction associated with these adjustments supports the view that overrides should be applied sparingly. While the extant study does not attempt to calculate or suggest what an optimal override rate might look like, future research should consider attempting to quantify an ideal override rate. Presently, all that exists is instrument developer suggestions that the override rate should be no more than 10% and perhaps not exceed 2% to 3% of assessments (Andrews et al., 2004; Casey et al., 2014).

CONCLUSION

This study shows that overrides of PCRA risk assessments nearly always lead to an upward reshuffling of supervisees into higher risk categories without any simultaneous improvement in prediction. In fact, overrides typically resulted in a degradation of the PCRA risk tool's predictive accuracy. This was particularly the case for supervisees with sex offender policy overrides, but it also held for those supervisees receiving discretionary or other policy overrides. These findings, along with prior research on supervision adjustments, suggest that overrides are associated with deteriorations in the predictive capacities of risk tools and that officers should use them sparingly and cautiously. Moreover, these findings suggest that future researchers consider attempting to calculate an optimal override rate for systems using actuarial tools.

APPENDIX

Equivalent Groups Generated by Propensity Score Matching—Any Override

	Panel A: Pre-matching group differences				Panel B: Post-matching group differences				
Matching covariates	Override supervisees (n = 27,480)	No override supervisees (n = 232,091)	T-statistic	p value	Override supervisees (n = 25,180)	No override supervisees (n = 25,180)	T-statistic	p value	% Bias reduction
PCRA domain scores									
Total score	6.35	7.54	–50.68*	.000	6.49	6.46	1.15	.248	97.0
Criminal history	3.65	4.50	–55.96*	.000	3.79	3.78	0.40	.686	99.0
Employment/education	1.06	1.26	–30.54*	.000	1.08	1.07	1.14	.256	95.1
Drugs/alcohol	0.28	0.46	–42.50*	.000	0.30	0.30	0.06	.950	99.8
Social networks	1.18	1.15	5.54*	.000	1.16	1.15	1.90	.058	51.6
Cognitions	0.17	0.15	7.50*	.000	0.16	0.15	1.25	.212	76.4
Gender (female = 1)	0.11	0.18	–29.90*	.000	0.12	0.12	–0.22	.826	99.1
Race									
White	0.54	0.33	70.38*	.000	0.51	0.51	0.66	.509	98.6
Black	0.25	0.36	–36.20*	.000	0.27	0.27	–0.65	.519	97.7
Hispanic, any race	0.14	0.25	–41.35*	.000	0.15	0.15	–1.14	.255	96.8
Other	0.07	0.06	6.27*	.000	0.07	0.07	1.44	.151	66.1
Age at assessment	43.09	39.41	49.41*	.000	42.56	42.58	–0.22	.829	99.4

Note. Nearest neighborhood matching with caliper 0.0001 was used. A total of 2,300 supervisees with overrides were lost through an inability to match with control group. In addition to matching on the above covariates, supervisees were matched across the 94 judicial districts (not shown in the table). The matching process resulted in balanced being achieved with p values exceeding the .001 threshold for each of the 94 judicial districts. The matching procedures are described in the Method section. An examination of the balancing for the sex offender and discretionary overrides is available on the longer paper uploaded to ResearchGate. PCRA = Post Conviction Risk Assessment.
*$p < .001$.

ORCID iD

Thomas H. Cohen https://orcid.org/0000-0002-0427-950X

REFERENCES

Administrative Office of the U.S. Courts. (2011). *An overview of the federal Post-Conviction Risk Assessment*.

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *Counseling Psychologist*, *34*(3), 341–382.

Andrews, D. A., & Bonta, J. (2010a). *The psychology of criminal conduct* (5th ed.). Mathew Bender.

Andrews, D. A., & Bonta, J. (2010b). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, *16*, 39–55.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*, 19–52.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *The Level of Service/Case Management Inventory*. Multi-Health Systems.

Borum, R., Otto, R., & Golding, S. (1993). Improving clinical judgment and decision making in forensic evaluation. *Journal of Psychiatry & Law*, *21*, 35–76.

Casey, P., Elek, J., Warren, R., Cheesman, F., Kleiman, M., & Ostrom, B. (2014). *Offender risk & needs assessment instruments: A primer for courts*. National Center for State Courts.

Chappell, A. T., Maggard, S. R., & Higgins, J. L. (2012). Exceptions to the rule? Exploring the use of overrides in detention risk assessment. *Youth Violence and Juvenile Justice*, *11*, 332–348.

Chouldechova, A. (2020). Transparency and simplicity in criminal risk assessment. *Harvard Data Science Review*, *2*(1). https://doi.org/10.1162/99608f92.b9343ee

Clem, C. (2003). *Topics in community corrections: Offender assessment*. National Institute of Corrections.

Cohen, T. H. (2018). Predicting sex offender recidivism: Using the federal Post Conviction Risk Assessment instrument to assess the likelihood of recidivism among federal sex offenders. *Journal of Empirical Legal Studies*, *15*(3), 456–481.

Cohen, T. H., Lowenkamp, C. T., & VanBenschoten, S. W. (2016). Examining changes in offender risk characteristics and recidivism outcomes: A research summary. *Federal Probation*, *80*(2), 57–65.

Cohen, T. H., Pendergast, B., & VanBenschoten, S. W. (2016). Examining overrides of risk classifications for offenders on federal supervision. *Federal Probation*, *80*(1), 12–21.

Cohen, T. H., & Spidell, M. (2016). How dangerous are they? An analysis of sex offenders under federal post-conviction supervision. *Federal Probation*, *80*(2), 21–32.

Cohen, T. H., & VanBenschoten, S. W. (2014). Does the risk of recidivism for supervised offenders improve over time: Examining changes in the dynamic risk characteristics for offenders under federal supervision. *Federal Probation*, *78*(2), 41–56.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674.

Gebo, E., Stracuzzi, N. F., & Hurst, V. (2006). Juvenile justice reform and the courtroom workgroup: Issues of perception and workload. *Journal of Criminal Justice*, *34*(4), 425–433.

Gottfredson, S. D., & Moriarty, L. J. (2006). Clinical versus actuarial judgments in criminal justice decisions: Should one replace the other? *Federal Probation*, *70*(2), 15–18.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical statistical controversy. *Psychology, Public Policy, and Law*, *2*, 293–323.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, *12*(1), 19–30.

Guay, J. P., & Parent, G. (2018). Broken legs, clinical overrides, and recidivism risk: An analysis of decisions to adjust risk levels with the LS/CMI. *Criminal Justice and Behavior*, *45*(1), 82–100.

Guide to Judiciary Policy. (2014). *Guide to judiciary policy: Volume 8, probation and pretrial services*. Administrative Office of the U.S. Courts.

Guo, S., & Fraser, M. W. (2014). *Propensity score analysis: Statistical methods and applications* (Vol. 11). SAGE.

Hamilton, M. (2015). Back to the future: The influence of criminal history on risk assessments. *Berkeley Journal of Criminal Law*, *20*(1), 75–133.

Harris, P. (2006). What community supervision officers need to know about actuarial risk assessment and clinical judgment. *Federal Probation*, *70*(2), 8–14.

Holsinger, A., Lurigio, A., & Latessa, E. (2001). Practitioners' guide to understanding the basis of assessing offender risk. *Federal Probation*, *65*(1), 46–51.

Lowenkamp, C. T., Holsinger, A. M., & Cohen, T. H. (2015). PCRA revisited: Testing the validity of the federal Post Conviction Risk Assessment (PCRA). *Psychological Services*, *12*(2), 149–157.

Lowenkamp, C. T., Johnson, J., VanBenschoten, S. W., Robinson, C., & Holsinger, A. M. (2013). The federal Post Conviction Risk Assessment (PCRA): A construction and validation study. *Psychological Services*, *10*(1), 87–96.

Luallen, J., Radakrishnan, S., & Rhodes, W. (2016). The predictive validity of the Post-Conviction Risk Assessment among federal offenders. *Criminal Justice and Behavior*, *43*(9), 1173–1189.

McCafferty, J. T. (2015). Professional discretion and the predictive validity of a juvenile risk assessment instrument: Exploring the overlooked principle of effective correctional classification. *Youth Violence and Juvenile Justice*, *15*(2), 103–118.

Meehl, P. E. (1957). When shall we use our heads instead of the formula? *Journal of Counseling Psychology*, *4*, 268–273.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. American Psychological Association.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38.

Schmidt, F., Sinclair, S. M., & Thomasdót, S. (2016). Predictive validity of the Youth Level of Service/Case Management Inventory with youth who have committed sexual and non-sexual offenses: The utility of professional override. *Criminal Justice and Behavior*, *43*(3), 413–430.

Schneider, A. L., Ervin, L., & Snyder-Joy, Z. (1996). Further exploration of the flight from discretion: The role of risk/need instruments in probation supervision decisions. *Journal of Criminal Justice*, *24*(2), 109–121.

Serin, R., Lowenkamp, C. T., Johnson, J., & Trevino, P. (2016). Using a multi-level risk assessment to inform case planning and risk management: Implications for officers. *Federal Probation*, *80*(2), 10–15.

Skeem, J., & Lowenkamp, C. T. (2016). Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology*, *54*(4), 680–712.

Vaswani, N., & Merone, L. (2014). Are there risks with risk assessment? A study of the predictive accuracy of the Youth Level of Service—Case Management Inventory with young offenders in Scotland. *British Journal of Social Work*, *44*, 2163–2181.

Wormith, J. S., Hogg, S., & Guzzo, L. (2012). The predictive validity of a general risk/needs assessment inventory on sexual offender recidivism and an exploration of the professional override. *Criminal Justice and Behavior*, *39*(2), 1511–1538.

Wormith, J. S., Hogg, S. M., & Guzzo, L. (2015). The predictive validity of the LS/CMI with aboriginal offenders in Canada. *Criminal Justice and Behavior*, *42*(5), 481–508.

Thomas H. Cohen received his PhD in criminal justice from the University of Rutgers School of Criminal Justice in 2003 and his JD from the University of Maryland School of Law in 1996. Currently, he is a social science analyst at the Administrative Office of the United States Courts (AOUSC), Probation and Pretrial Services Office. His work includes analyzing risk assessment data at the post-conviction and pretrial levels and authoring reports on how the AOUSC integrates the risk principle into its operational practices. His recent research has appeared in *Criminology and Public Policy, Criminal Justice Policy Review, Federal Probation, Journal of Empirical Legal Studies*, and *Psychological Services*. Moreover, he has authored several technical reports on criminal court case processing at the state and federal levels through his prior work at the Bureau of Justice Statistics.

Christopher T. Lowenkamp received his PhD in criminal justice from the University of Cincinnati. He has served as the director of the Center for Criminal Justice Research, the associate director of The Corrections Institute at the University of Cincinnati, and research professor at the University of Cincinnati. He is currently a social science analyst for the Administrative Office of the United States Courts (AOUSC). He has co-authored more than 65 articles and book chapters, some of which are published in top-tier academic journals.

Kristin Bechtel received her PhD in criminal justice from the University of Cincinnati. She serves as Director of Criminal Justice Research at Arnold Ventures and oversees an extensive pretrial justice research portfolio. Her work focuses on the evaluation of policies and interventions aimed to reduce mass incarceration and improve outcomes for justice-involved persons and their communities.

Anthony W. Flores earned his PhD in criminal justice from the University of Cincinnati and is currently an associate professor in criminal justice at California State University, Bakersfield. His research interests include offender risk and need assessment, effective rehabilitative programming strategies, and the prediction of recidivism.