



**Pennsylvania
Commission
on
Sentencing**

State College Office:
009 Brumbaugh Hall
Penn State
University Park Campus

Mail:
PO Box 1200
State College, PA
16804-1200

Phone:
814.863.2797

Fax:
814.863.2129

URL:
<http://pasentencing.us>

Risk/Needs Assessment Project

Interim Report 5

Developing Categories of Risk

2012



Risk/Needs Assessment Project

Interim Report 5 Developing Categories of Risk

Highlights

This report presents findings from a series of analyses examining the accuracy of risk prediction for various risk classification groupings. There were five major findings:

- About 52% of the total development sample had a re-arrest after three years.
- The factors that best predicted recidivism in our analyses of ‘solitary conviction offenders’ were found to also best predict recidivism for ‘multiple conviction offenders’. Further, the number of current convictions did not impact recidivism.
- There were eight factors found to predict recidivism, which resulted in a risk score ranging from 0-14.
- We grouped the risk scores [0-14] into two categories of risk [low/high] and three categories of risk [low/medium/high] in order to examine the accuracy and error rates for various groupings. The overall prediction accuracy of the various models we created ranged from 49% to 65%, depending upon how the risk scores were grouped together.
- More importantly, there are trade-offs in the prediction accuracy of low risk and high risk, depending upon how the risk scores are classified into groups. The lower the cutoff point for determining high risk, the better the prediction accuracy for high risk [which ranged from 53% to 80%], but the worse the prediction accuracy for low risk [which ranged from 48% to 86%]. Alternatively, the higher the cutoff point for determining high risk, the worse the prediction accuracy for high risk, but the better the prediction accuracy for low risk.

In the final determination of the risk classification groups there are two policy issues to be considered:

- *What is the potential impact of risk categories on sentencing practices?* It is important to know the impact of the potential risk score categories on sentencing practices. For example, how many offenders would be projected to receive a harsher sentence compared to how many people would be projected to receive a more lenient sentence? What would be the overall net impact for the various options of risk classification categories?
- *False positive/false negative error tradeoffs.* Is it better to err on the side of over predicting arrest [which potentially could result in correctional overcrowding] or under predicting arrest [which potentially could result in more crime]?



Background.

The correctional reform legislation enacted in 2008 [Acts 81, 82, 83 and 84 of 2008] requires the Commission to develop and adopt new guidelines for parole (county and state) and re-parole, as well as for re-sentencing following revocation of probation, CIP, and SIP. In developing guidelines for parole, Act 81 of 2008 mandates that the guidelines incorporate validated risk assessment tools, and take into account available research relating to the risk of recidivism. Additionally, Act 95 of 2010, which was signed by Governor Rendell on October 27, 2010, mandated the Commission to develop a risk assessment instrument to assist the court at sentencing.

To address these new mandates, the Commission has undertaken a recidivism study to determine what factors best predict recidivism for various types of offenders. This study is utilizing primarily the criminal justice factors, and the limited demographic factors that are currently collected by the Commission on Sentencing and State Police, and that are therefore readily available in the field. Future analyses will expand the study to determine the impact that other, less available, factors [e.g., drug use, employment status, education] have on the risk of recidivism.

Recidivism Study

Sample. In an effort to learn more about what risk factors best predict recidivism for certain types of offenders, we undertook a recidivism study of Level 3 and Level 4 offenders, along with a limited number of Level 5 offenders [OGS/PRS of 9/0; 9/1; 9/2]. These levels were chosen for three reasons: 1) the offenses at these levels encompass a wide variety of offense seriousness [OGS ranging from 2 to 8 depending upon prior record]; 2) the sentence recommendations provide for a variety of sentence types [prison, SIP, BC, jail, probation, IP]; and 3) we could use a sample sentenced during 2004, 2005, and 2006 [SGS Web data years], which allowed for a three year tracking period for most offenders.

Measuring Recidivism. We used re-arrest, and re-incarceration on a technical violation for offenders sentenced to state prison, as our measures of recidivism. We obtained arrest information from the criminal history records maintained by the State Police. In determining exposure time, we used date of sentence for probation cases, expiration of minimum sentence for county jail sentences,¹ and date of release for state prison sentences. The Department of Corrections provided the date of release, as well as information on technical violations that resulted in return to prison, which we took into account for the recidivism of offenders sentenced to prison.

Study Sample: Development and Validation Samples. We decided to examine DUI offenders separately because they differ from non-DUI offenders on several dimensions. Thus, our total sample consisted of 41,563 non-DUI offenders.² Following best practices in risk assessment research, we then randomly split

¹ For county jail sentences, date of release was unavailable. Since the judge has paroling authority in these cases, offenders can be released prior to the expiration of the minimum. Findings from another study the Commission has underway indicate that about a third of the offenders are released prior to their minimum, about a third at the expiration of their minimum, and about a third post minimum sentence. Further, most offenders in our study had served their maximum sentence for the two and three year recidivism analysis. Thus, we decided that expiration of minimum sentence was the best measure to use to determine exposure time.

² The sample began with 58,696 offenders based upon separate judicial proceedings. Cases were removed due to SID not being found [n=124]; lack of match with state police criminal history records [n=886]; 'sample offense' not located in criminal history records [n=3,677]; cases entered as separate judicial proceeding but should have been entered as single judicial proceeding [n=3,075]; offender appeared twice in same year [6,417]; offender had less than one year follow-up [n=1,409]; no match with



the sample into two sub-samples; one, a development sample to be used for the development of the risk assessment tool, the other, a validation sample to test the predictive capability of the tool. This split resulted in a sample of 20,812 for the development sample.

Previous Interim Reports.

This is the fifth report in a series of interim reports on the progress of the project. The four previous reports that have been provided on the project covered the following topics:

Interim Report 1: Factors Considered by Current Risk Instruments. Stage I of the project, which was started during the summer of 2010, was directed toward obtaining information on the current utilization of risk assessment instruments in other jurisdictions, the identification of risk factors used by other instruments, and the availability of information contained in pre-sentence investigations in a sample of 25 counties. A report on the findings from that stage was provided to the Commission in October 2010.

Interim Report 2: Recidivism Rates by Various Factors. Stage II of the project was the initial recidivism study utilizing primarily criminal justice factors, and limited demographic factors, to determine what factors best predict recidivism for various types of offenders. Report 2 was provided to the Commission in June 2011 and provided recidivism rates with respect to:

- Overall recidivism rates after one year, two year, and three year tracking periods
- Recidivism by OGS and PRS
- Recidivism by guideline cell
- Recidivism by sentence type
- Recidivism by offense type
- Recidivism by specific offenses within each guideline cell

This report presented the findings from a series of analyses conducted to determine which factors are the best predictors of recidivism for offenders convicted of various types of offenses.

Interim Report 3: Factors that Predict Recidivism for Various Types of Offenders. The third Interim Report, which was provided to the Commission in December 2011, presented the findings from a series of analyses conducted to determine which factors are the best predictors of recidivism for offenders convicted of personal, property, drug, and firearms offenses. There were four major findings:

- While offenders do commit a variety of offenses, there is a tendency for offenders to specialize in the type of crime they commit.
- The most consistent predictors of recidivism were age and number of prior arrests. Younger offenders, and offenders with a greater number of prior arrests, were more likely to recidivate.
- Less consistent predictors of recidivism were race, gender, county, and offense gravity score. Black offenders were more likely than white offenders to recidivate; males were more likely than females to recidivate, and offenders from Philadelphia were more likely to recidivate, especially compared to offenders from rural counties. Offenders with lower offense gravity scores were more likely to recidivate.

DOC data [n=346]; offender still in prison [n=913]; offender transferred to state hospital [n=53]; offender escaped [n=135] , or offender had died during study period [n=98].



- Prior convictions and type of sentence were not found to predict recidivism.

Interim Report 4: Development of Risk Assessment Scale: Testing Three Classification Methods. The factors that were found to be predictive of recidivism in our study were used in the development of the risk assessment instrument. We examined three different classification methods to determine whether one method was better than the others, balanced with the ease of application and understanding:

- 1) *Burgess Method.* In this method, equal weight is given to each item that was found to be predictive in multivariate analyses.
- 2) *Weighted Burgess Method.* For this method, we gave some items more weight than others, based upon how strongly correlated the factors were to recidivism in multivariate analyses.
- 2) *Predictive Attribute Analysis.* For this method, risk classification groups were determined through a hierarchical clustering process.

We found that there was no significant difference in the predictive ability of the three methods. Thus, the Burgess Method was deemed to be the best classification method due to its ease of use and the ease to which it could be understood and explained.

Current Report: Determining Categories of Risk.

In this report, we discuss the analysis we did to determine how to categorize the risk assessment score into categories of risk. In the development of the risk assessment scale, we had included only those offenders with a single conviction [n=9,536], which enabled us to better examine the influence of offense type. For this subsequent analysis, we are including all of the offenders in the development sample. There were 8,725 offenders with multiple convictions, so that the total development sample is 18,261 offenders.³ Before we could do the risk classification analyses, we needed to first run our previous recidivism analysis on the entire sample to see whether the findings changed or stayed the same. Our analyses used offenders who had a three-year follow-up period. These findings are discussed below.

Sample Description. Table 1 shows the description of the full sample, which is very similar to the sample of solitary offenders. Most of the offenders were male [86%], from an urban county [83%], and had a mean age of 31. Almost half, 46%, were Black; 44% were white, and about 9% were Hispanic. The average Offense Gravity Score was 5 [based on a scale of 1- 9 used for this sample], with the largest number of offenders being convicted of a drug offense [42%], followed by property [28%], personal [18%], firearms [5%], and other [7%] offenses. The number of current convictions ranged from 1 to 96. About 48% of the sample had more than one current conviction. However, 26% had two convictions, 11% had three convictions, and 11% had more than three convictions. The majority [85%] of offenders had at least one prior arrest, and had a previous arrest for a personal [60%], property [60%], and/or drug [61%] offense. Most of the offenders had prior convictions [70%], with a mean PRS of 2 [on a scale of 0 to 6]. The most common sentence imposed was jail [56%], with the remaining offenders receiving prison

³ The original development sample was 20,812. For the analyses in this report, we had removed those offenders who had technical violations only, primarily because we did not have this information on county jail offenders and it kept the analysis cleaner. This removed 691 offenders from the analysis. An additional 1,830 offenders were removed because they did not have a three year follow-up time period.



Pennsylvania Commission on Sentencing

Table 1. Descriptive statistics for full development sample of offenders (N = 18,261).¹

	N	%	% 3-year recidivism		N	%	% 3-year recidivism
Gender***				Type of prior arrest(s)			
Male	15,642	85.7	53.9	Prior personal/sex arrest(s)***			
Female	2,619	14.3	41.5	Yes	9,562	52.4	59.8
	18,261	100.0	52.1	No	8,699	47.6	43.8
					18,261	100.0	52.1
County***				Prior property arrest(s)***			
Philadelphia	4,113	22.5	60.3	Yes	12,219	66.9	59.0
Allegheny	2,648	14.5	60.8	No	6,042	33.1	38.2
Other urban	8,474	46.4	49.7		18,261	100.0	52.1
Rural	3,026	16.6	40.4				
	18,261	100.0	52.1	Prior drug arrest(s)***			
Age***				Yes	9,834	53.9	60.8
< 21	3,348	18.3	62.0	No	8,427	46.2	42.0
21 - 24	3,266	17.9	56.6		18,261	100.0	52.1
25-29	2,959	16.2	52.2	Prior firearms/weapons arrest(s)***			
30-34	2,279	12.5	51.0	Yes	3,179	17.4	64.4
35-39	2,176	11.9	50.5	No	15,082	82.6	49.6
40-44	1,975	10.8	47.5		18,261	100.0	52.1
45-49	1,261	6.9	42.4	Prior other arrest(s)***			
> 50	997	5.5	31.7	Yes	12,585	68.9	58.7
	18,261	100.0	52.1	No	5,676	31.1	37.6
Mean	30.97				18,261	100.0	52.1
Race***				PRS***			
White	8,115	44.4	44.8	0	5,668	31.0	43.2
Black	8,336	45.7	59.7	1	2,137	11.7	50.5
Hispanic	1,606	8.8	50.5	2	2,128	11.7	55.1
Other	204	1.1	49.0	3	1,256	6.9	55.6
	18,261	100.0	52.1	4	2,194	12.0	56.7
OGS***				5	3,879	21.2	59.6
1	177	1.0	54.2	RFEL	999	5.5	56.7
2	707	3.9	59.3		18,261	100.0	52.1
3	3,529	19.3	59.0	Mean	2.43		
4	439	2.4	50.8	Multiple charges			
5	3,237	17.7	55.3	Yes	8,725	47.8	52.2
6	4,918	26.9	50.8	No	9,536	52.2	52.1
7	3,605	19.7	48.2		18,261	100.0	52.1
8	1,186	6.5	36.3	Mean	2.05		
9	463	2.5	52.9	Type of sentence***			
	18,261	100.0	52.1	Prison	2,241	12.3	52.2
Mean	5.39			SIP	12	0.1	33.3
Current offense type (most serious)***				Jail	10,290	56.4	53.1
Personal-- Felony	1,814	9.9	7.5	CIP	1,921	10.5	49.5
Personal-- Misdemeanor	1,179	6.5	48.4	Probation	3,556	19.5	50.7
Sex offense-- Felony	360	2.0	36.7	Other	241	1.3	53.5
Sex offense-- Misdemeanor	98	0.5	49.0		18,261	100.0	52.1
Drug-- Felony (PWID)	6,478	35.5	50.9	Recidivism			
Drug-- Misdemeanor (Possession)	1,130	6.2	59.7	One Year			
Burglary	1,060	5.8	53.4	Yes	5,006	27.4	
Other property offense	4,079	22.3	55.7	No	13,255	72.6	
Firearms/other weapons	930	5.1	57.7	Two Year			
Other offense	1,133	6.2	46.7	Yes	7,822	42.8	
	18,261	100.0	52.1	No	10,439	57.2	
Prior arrest(s)***				Three Year			
Yes	15,525	85.0	56.3	Yes	9,521	52.1	
No	2,736	15.0	28.5	No	8,740	47.9	
	18,261	100.0					
Total prior arrests***							
0	2,736	15.0	28.5				
1	2,241	12.3	40.9				
2-4	5,213	28.5	52.4				
5-12	6,049	33.1	61.0				
13+	1,403	7.7	69.4				
	18,261	100.0	52.1				
Mean	5.60						

* p < .05 ** p < .01 *** p < .001 (Chi-square/ T-test for variable and 3-year recidivism)

¹ Excludes offenders who did not have a three-year follow-up (N = 1,830) and offenders whose first failure was a technical violation (N = 691).



[12%], probation [19%], and county intermediate punishment [11%]. About 52% of the sample were re-arrested within three years

Testing the recidivism model with entire development sample. Before conducting the analysis for grouping of risk classification categories, we first conducted the recidivism analysis with the full development sample to determine whether the findings of recidivism predictors were the same with all offenders as with the solitary conviction sample. We then included the variable ‘number of current convictions’ to determine whether this factor altered the findings or was a significant factor itself. We found that the findings remained the same with the full development sample, and that the inclusion of ‘number of current convictions’ did not alter the findings and was a significant factor in predicting recidivism [see Table 1A in Appendix]. Thus, our scoring of factors used in the Burgess Risk classification model remained the same. Table 2 shows the factors included in the risk scale, which resulted in a risk score ranging from 0-14.

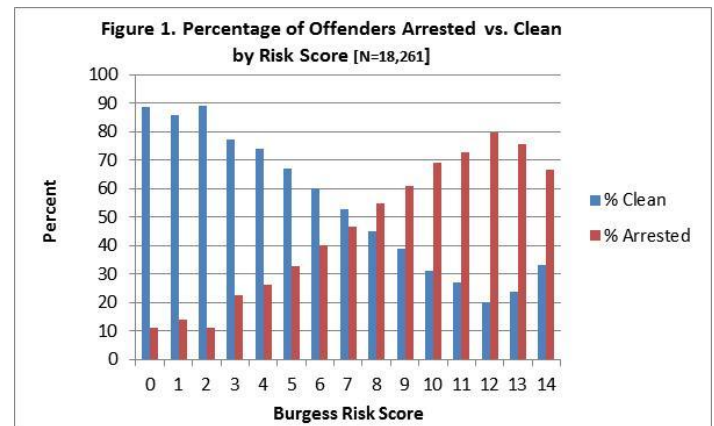
Table 2. Burgess Risk Scale for Development Sample Offenders (N = 18,261).	
FACTOR	Score *
Gender	
Female	0
Male	1
Age	
Less than 24	3
24-29	2
30-49	1
50+	0
County	
Rural counties	0
Smaller, urban counties	1
Allegheny and Philadelphia Counties	2
Total number of prior arrests	
0	0
1	1
2 to 4	2
12-May	3
13+	4
Prior property arrests	
No	0
Yes	1
Prior drug arrests	
No	0
Yes	1
Property offender	
No	0
Yes	1
Offense gravity score (OGS)	
4+	0
1 to 3	1
* Total Possible Range is 0 to 14.	



Table 3 shows the number of offenders by risk score [0-14] and the percentage of offenders who were arrested by risk score. Small percentages of offenders had either very low risk scores [0-4; 12%] or very high risk scores [11-14; 11%]. More than three-quarters [77%] had risk scores in the middle range [5-10]. Figure 5 shows a positive relationship between the risk score and recidivism; the higher the risk score, the higher the rate of re-arrest.

Table 3. Three-year Recidivism by Burgess Risk Score Among Development Sample Offenders (N = 18,261).

Burgess Risk Score	Number	Percent	% Not Arrested	% Arrested	Number Not Arrested	Number Arrested
0	9	0.1	88.9	11.1	8	1
1	78	0.4	85.9	14.1	67	11
2	300	1.6	89.0	11.0	267	33
3	677	3.7	77.4	22.6	522	155
4	1,121	6.1	74.0	26.3	821	300
5	1,545	8.5	67.0	32.8	1,030	515
6	1,702	9.3	60.0	40.3	1,004	698
7	2,118	11.6	53.0	46.6	1,114	1004
8	2,829	15.5	45.0	54.7	1,262	1567
9	3,101	17.0	39.0	61.0	1,188	1913
10	2,718	14.9	31.0	69.0	814	1904
11	1,598	8.8	27.0	73.0	408	1190
12	429	2.4	20.1	80.0	84	345
13	33	0.2	24.0	75.8	8	25
14	3	0.0	33.3	66.7	1	2
Total	18,261	100.0	47.9	52.1	8598	9663



Developing Categories of risk

After the scale was developed, the next step was to determine how to group offenders into categories of risk. One option was to define two categories of recidivism risk: low risk and high risk. However, if the risk category was to be incorporated into the guidelines structure, three risk categories may be deemed more appropriate: low, medium, and high. With three risk categories, the guidelines might recommend that a medium risk score be a factor that warrants consideration of a standard range sentence, while a high risk score might be considered an aggravating factor, and a low risk score a mitigating factor.

The analyses for examining potential categories of risk included comparing the prediction accuracy and error rates among the possible risk categories. We conducted these analyses when using both two categories of risk and three categories of risk.

Two categories. We first examined the various options for grouping offenders into two risk categories. High risk was defined as forecasting a future arrest and low risk was defined as forecasting no future arrest. Table 4 shows the comparison of accuracy rates for the forecasts of high risk versus low risk offenders for all of the various cut-off points on the risk scale. The overall accuracy of the predictions ranges from 49% to 65%, depending on the cut-off point. However, the accuracy of the forecast varies for the high and low risk groups depending on where the cut-off point for the two risk groups is set. For example, Table 5 shows that when the cut-off score for high risk is 8 [i.e., having a risk score of 8 through 14], 59% of the offenders [n=10,711] are forecasted to be high risk and 41% of the offenders [n=7,550] are forecasted to be low risk. Of those forecasted to be high risk, 65% [n=6,946] turned out to be high



risk [i.e., arrested]. Of those forecasted to be low risk, 64% [n=4,833] turned out to be low risk [i.e., were not arrested.]

Table 4. Comparison of Different Cutoff Points and Error Rates for Original and New Burgess Risk Scales, Among Development Sample Offenders. Outcome Defined as Arrest vs. No Arrest (N = 18,261).

	% Forecast high risk	% Forecast low risk	Total percent of forecasts that were accurate ¹	Of those forecast to be high risk, % that actually were	Of those forecast to be low risk, % that actually were	Of those actually High Risk, % forecasted accurately	Of those actually Low Risk, % forecasted accurately	Overall false positive rate ²	Overall false negative rate ³	False positive/negatives ratio, ⁴	False negative/false positive ratio ⁵
Variations on 2 group combinations											
2 groups: (L: 0-1 H: 2-14)	99.5%	0.5%	53.3%	53.1%	86.2%	99.9%	0.9%	46.9%	13.8%	710.25	0.00
2 groups: (L: 0-2 H: 3-14)	97.9%	2.1%	54.5%	53.8%	88.4%	99.5%	4.0%	46.2%	11.6%	183.47	0.01
2 groups: (L: 0-3 H: 4-14)	94.2%	5.8%	56.6%	55.0%	81.2%	97.9%	10.0%	45.0%	18.8%	38.67	0.03
2 groups: (L: 0-4 H: 5-14)	88.0%	12.0%	59.4%	57.0%	77.1%	94.8%	19.6%	43.0%	22.9%	13.83	0.07
2 groups: (L: 0-5 H: 6-14)	79.6%	20.4%	62.2%	59.5%	72.8%	89.5%	31.6%	40.5%	27.2%	5.80	0.17
2 groups: (L: 0-6 H: 7-14)	70.3%	29.7%	63.9%	62.0%	68.5%	82.3%	43.3%	38.0%	31.5%	2.85	0.35
2 groups: (L: 0-7 H: 8-14)	58.7%	41.3%	64.5%	64.8%	64.0%	71.9%	56.2%	35.2%	36.0%	1.39	0.72
2 groups: (L: 0-8 H: 9-14)	43.2%	56.8%	62.8%	68.2%	58.7%	55.7%	70.9%	31.8%	41.3%	0.58	1.71
2 groups: (L: 0-9 H: 10-14)	26.2%	73.8%	58.9%	72.5%	54.0%	35.9%	84.7%	27.5%	46.0%	0.21	4.71
2 groups: (L: 0-10 H: 11-14)	11.3%	88.7%	52.9%	75.7%	50.0%	16.2%	94.2%	24.3%	50.0%	0.06	16.17
2 groups: (L: 0-11 H: 12-14)	2.5%	97.5%	48.6%	80.0%	47.8%	3.8%	98.9%	20.0%	52.2%	0.01	99.90

Note. Low risk means that we predicted no arrest. High risk means that we predicted any arrest. Highlighted row indicates where the error rates are about the same for low risk and high risk groups [L: 0-7; H: 8-14].

¹ This figure calculated by taking: (low risk offenders who were not arrested + high risk offenders who were arrested)/(total offenders)

² False positives are high risk offenders who were not arrested. This figure was calculated by taking: (high risk offenders who were not arrested / total high risk offenders)

³ False negative are low risk offenders who were arrested. This figure was calculated by taking: (low risk offenders who were arrested / total low risk offenders).

⁴ This figure was calculated by taking: (high risk offenders who were not arrested / low risk offenders who were arrested).

⁵ This figure is the inverse of the false positive/false negatives for high risk offenders. It was calculated by taking: (low risk offenders who were arrested / high risk offenders who were not arrested).

Table 5. Actual and Predicted Risk When Cut-off Score is 8 on Burgess Risk Scale.(N = 18,261).

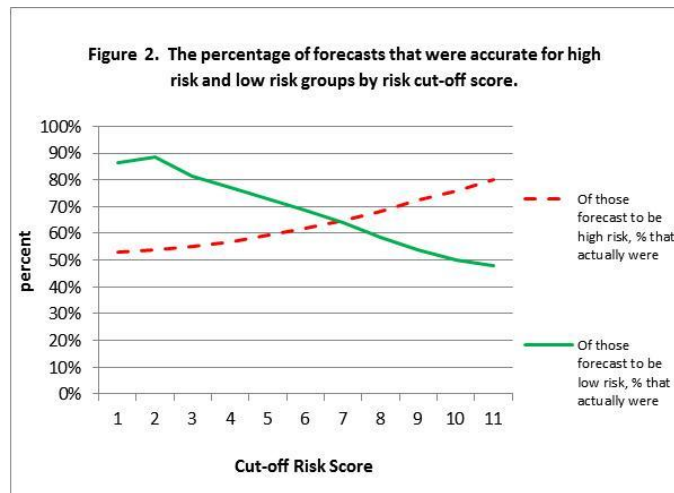
Predicted	Actual		Total
	No Arrest	Arrest	
Low risk (No arrest/0-7)	4833	2717	7550
High risk (Arrest/8-14)	3765	6946	10711
Total	8598	9663	18261

Boxes shaded in green were predicted correctly.

Boxes shaded in red were predicted incorrectly.



Figure 2 shows that the higher the cut-off point is set, the higher the accuracy for the high risk group; but the accuracy for the low risk group gets corresponding lower. The lower the cut-off point is set, the higher the accuracy for the low risk group; but the accuracy of the high risk group gets corresponding lower. Figure 2 also shows that when the cut-off is about 8 the accuracy for both high and low risk groups is about the same. Another way to think about it is that the percentages of false positives [high risk offenders who were not arrested] and false negatives [low risk offenders who were arrested] were about equal [35% vs. 36%].



In order to gain a better understanding of the trade-offs in the accuracy of the forecasts for the high risk and low risk groups, we can look at the false positive/false negative ratio from Table 4. This rate indicates how often there is a false prediction of arrest. Alternatively, we can also look at the false negative/false positive ratio, which indicates how often there is a false prediction of no arrest. Table 6 shows the false positive/false negative ratios and the false negative/false positive ratios for the different cutoff points that can be used to define a high risk group. A low cutoff score includes more people forecasted to be arrested, and thus, results in an over prediction of arrest. For example, at the lowest category [defines a risk score of 2-14 to be high risk] there were 710 people falsely predicted to recidivate for every one person accurately predicted to recidivate. On the other hand, a high cut off score includes fewer people defined as high risk, which results in an under prediction of arrest. For example, in the highest

Table 6. The false positive/false negative ratio for risk group categories.

Risk group category	Overprediction of arrest	Underprediction of arrest
	for every person correctly predicted to recidivate the number who did not recidivate	
2-14	710.25	
3-14	183.47	
4-14	38.67	
5-14	13.83	
6-14	5.80	
7-14	2.85	
8-14)	1.39	
		for every person correctly predicted to not recidivate the number who did recidivate
9-14		1.71
10-14		4.71
11-14		16.17
12-14		99.90



category [defines a risk score of 12-14 as high risk], the false negative/false positive rate is 99.9 indicating that there were 100 people falsely predicted to not recidivate for every one person accurately predicted to not recidivate. As the cut-off point moves toward to middle, the trade-offs become less extreme. As indicated earlier, there is basically a one to one trade-off in the error rates when high risk is defined as anyone with a risk score of 8-14. The primary policy issue is whether it is better to risk over prediction of arrest [potential increase in incarceration] or under prediction of arrest [potential increase of crime].

Three categories. We next examined how the cut-off points could be established if there were three categories of risk: high, medium, and low. This is more of a challenge as the prediction is no longer mutually exclusive as it was with the two groups [i.e., high risk means arrest; low risk means no arrest]. That is, the medium risk group has no clear independent predictor, and thus, a false positive and false negative rate cannot be determined for the three groups of high, medium, and low risk. The method we employed to address this problem was as follows. First, we divided the offenders into two groups of risk to examine the accuracy of prediction for various cut off points for low risk when high risk was held constant. Second, we divided the offenders into two groups of risk to examine the accuracy of prediction for various cut off points of high risk when low risk was held constant. After determining the error rates and accuracy of the high risk and low risk categories, by default, the medium risk group was the ‘leftover’ group of offenders.

Table 7 shows the accuracy of prediction for the various cut-off points for low risk when the high risk group was defined with three different ranges of risk scores: 10-14, 11-14, and 12-14. In our previous analysis of two categories, we found that for all three of these categories, over 70% of the offenders predicted to be high risk actually were high risk [see Table 4]. Thus, these categories seemed reasonable to set as high risk categories for this analysis. Table 8 compares the false positive/false negative ratio for these three risk groups. The non-shaded area shows where there is basically a one to one tradeoff in prediction. For every offender falsely predicted to have an arrest, one offender was falsely predicted to not have an arrest. The light grey shaded area shows the error is more in the direction of over predicting arrest [ratio above 1] and the dark shaded area shows where the error is more likely to be in the direction of under predicting arrest [ratio below 1]. This pattern indicates that to obtain a one to one trade off in the error rate, the low risk category needs to have a different cutoff point depending upon how the high risk category is defined.



Pennsylvania Commission on Sentencing

Table 7. Comparison of Different Cutoff Points and Error Rates for Original and New Burgess Risk Scales, Among Development Sample Offenders. Outcome Defined as Arrest vs. No Arrest (N = 18,261).

	% Forecast high risk	% Forecast medium risk	% Forecast low risk	Total percent of forecasts that were accurate ¹	Of those forecast to be high risk, % that actually were	Of those forecast to be low risk, % that actually were	Of those actually High Risk, % forecasted accurately	Of those actually Low Risk, % forecasted accurately	Overall false positive rate ²	Overall false negative rate ³	Overall false positives/f false negatives	False positive/ false negatives, High risk ⁴	False positive/ false negatives, Low risk ⁵
Variations on low risk cutoff value (High risk: 10-14)													
3 groups: (L: 0-1 H: 10-14)	26.2%	73.3%	0.5%	72.7%	72.5%	86.2%	99.7%	5.4%	27.5%	13.8%	109.58	109.58	0.01
3 groups: (L: 0-2 H: 10-14)	26.2%	71.7%	2.1%	73.7%	72.5%	88.4%	98.7%	20.6%	27.5%	11.6%	29.22	29.22	0.03
3 groups: (L: 0-3 H: 10-14)	26.2%	68.0%	5.8%	74.1%	72.5%	13.5%	94.5%	39.7%	27.5%	18.8%	6.58	6.58	0.15
3 groups: (L: 0-4 H: 10-14)	26.2%	61.9%	12.0%	73.9%	72.5%	77.1%	87.4%	56.2%	27.5%	22.9%	2.63	2.63	0.38
3 groups: (L: 0-5 H: 10-14)	26.2%	53.4%	20.4%	72.6%	72.5%	72.8%	77.3%	67.4%	27.5%	27.2%	1.30	1.30	0.77
3 groups: (L: 0-6 H: 10-14)	26.2%	44.1%	29.7%	70.4%	72.5%	68.5%	66.9%	73.9%	27.5%	31.5%	0.77	0.77	1.30
3 groups: (L: 0-7 H: 10-14)	26.2%	32.5%	41.3%	67.3%	72.5%	64.0%	56.1%	78.6%	27.5%	36.0%	0.48	0.48	2.07
3 groups: (L: 0-8 H: 10-14)	26.2%	17.0%	56.8%	63.1%	72.5%	58.7%	44.7%	82.3%	27.5%	41.3%	0.31	0.31	3.26
3 groups: (L: 0-9 H: 10-14)	26.2%	0.0%	73.8%	58.9%	72.5%	54.0%	35.9%	84.7%	27.5%	46.0%	0.21	0.21	4.71
Variations on low risk cutoff value (High risk: 11-14)													
3 groups: (L: 0-1 H: 11-14)	11.3%	88.2%	0.5%	76.1%	75.7%	86.2%	99.2%	13.0%	24.3%	13.8%	41.75	41.75	0.02
3 groups: (L: 0-2 H: 11-14)	11.3%	86.6%	2.1%	77.7%	75.7%	88.4%	97.2%	40.6%	24.3%	11.6%	11.13	11.13	0.09
3 groups: (L: 0-3 H: 11-14)	11.3%	82.9%	5.8%	77.6%	75.7%	81.2%	88.6%	63.3%	24.3%	18.8%	2.51	2.51	0.40
3 groups: (L: 0-4 H: 11-14)	11.3%	76.7%	12.0%	76.4%	75.7%	77.1%	75.8%	77.1%	24.3%	22.9%	1.00	1.00	1.00
3 groups: (L: 0-5 H: 11-14)	11.3%	68.3%	20.4%	73.8%	75.7%	72.8%	60.6%	84.4%	24.3%	27.2%	0.49	0.49	2.03
3 groups: (L: 0-6 H: 11-14)	11.3%	59.0%	29.7%	70.5%	75.7%	68.5%	47.7%	88.1%	24.3%	31.5%	0.29	0.29	3.42
3 groups: (L: 0-7 H: 11-14)	11.3%	47.4%	41.3%	66.5%	75.7%	64.0%	36.5%	90.6%	24.3%	36.0%	0.18	0.18	5.42
3 groups: (L: 0-8 H: 11-14)	11.3%	31.9%	56.8%	61.5%	75.7%	58.7%	26.7%	92.4%	24.3%	41.3%	0.12	0.12	8.55
3 groups: (L: 0-9 H: 11-14)	11.3%	14.9%	73.8%	56.9%	75.7%	54.0%	20.1%	93.6%	24.3%	39.9%	0.08	0.08	12.37
3 groups: (L: 0-10 H: 11-14)	11.3%	0.0%	88.7%	52.9%	75.7%	50.0%	16.2%	94.2%	24.3%	50.0%	0.06	0.06	16.17
Variations on low risk cutoff value (High risk: 12-14)													
3 groups: (L: 0-1 H: 12-14)	2.5%	97.0%	0.5%	81.0%	80.0%	86.2%	96.9%	44.6%	20.0%	13.8%	7.75	7.75	0.13
3 groups: (L: 0-2 H: 12-14)	2.5%	95.3%	2.1%	83.8%	80.0%	88.4%	89.2%	78.6%	20.0%	11.6%	2.07	2.07	0.48
3 groups: (L: 0-3 H: 12-14)	2.5%	91.6%	5.8%	80.8%	80.0%	81.2%	65.0%	90.3%	20.0%	18.8%	0.47	0.47	2.15
3 groups: (L: 0-4 H: 12-14)	2.5%	85.5%	12.0%	77.6%	80.0%	77.1%	42.7%	94.8%	20.0%	22.9%	0.19	0.19	5.38
3 groups: (L: 0-5 H: 12-14)	2.5%	77.0%	20.4%	73.6%	80.0%	72.8%	26.8%	96.7%	20.0%	27.2%	0.09	0.09	10.91
3 groups: (L: 0-6 H: 12-14)	2.5%	67.7%	29.7%	69.4%	80.0%	68.5%	17.8%	97.6%	20.0%	31.5%	0.05	0.05	18.42
3 groups: (L: 0-7 H: 12-14)	2.5%	56.1%	41.3%	64.9%	80.0%	64.0%	12.0%	98.1%	20.0%	36.0%	0.03	0.03	29.22
3 groups: (L: 0-8 H: 12-14)	2.5%	40.6%	56.8%	59.6%	80.0%	58.7%	8.0%	98.5%	20.0%	41.3%	0.02	0.02	46.06
3 groups: (L: 0-9 H: 12-14)	2.5%	23.6%	73.8%	54.9%	80.0%	54.0%	5.7%	98.7%	20.0%	46.0%	0.02	0.02	66.63
3 groups: (L: 0-10 H: 12-14)	2.5%	8.8%	88.7%	50.8%	80.0%	50.0%	4.4%	98.9%	20.0%	50.0%	0.01	0.01	87.11
3 groups: (L: 0-11 H: 12-14)	2.5%	0.0%	97.5%	48.6%	80.0%	47.8%	3.8%	98.9%	20.0%	52.2%	0.01	0.01	99.90

Note. Low risk means that we predicted no arrest. Medium and high risk mean that we predicted any arrest. Only totals for low and high risk offenders were used to predict forecast accuracies.

¹ This figure calculated by taking: (low risk offenders who were not arrested + high risk offenders who were arrested) / (total high and low risk offenders)

² False positives are high risk offenders who were not arrested. This figure was calculated by taking: (high risk offenders who were not arrested / total high risk offenders)

³ False negative are low risk offenders who were arrested. This figure was calculated by taking: (low risk offenders who were arrested / total low risk offenders)

⁴ This figure was calculated by taking: (high risk offenders who were not arrested / low risk offenders who were arrested)

⁵ This figure is the inverse of the false positive/false negatives for high risk offenders. It was calculated by taking: (low risk offenders who were arrested / high risk offenders who were not arrested)

Table 8. False positive to false negative ratios for High Risk Group

	High Risk Group Categories		
	10 to 14	11 to 14	12 to 14
Cut-off for low risk			
1	109.58	41.75	7.75
2	29.22	11.13	2.07
3	6.58	2.51	0.47
4	2.63	1.00	0.19
5	1.30	0.49	0.09
6	0.77	0.29	0.05
7	0.48	0.18	0.03
8	0.31	0.12	0.02
9	0.21	0.08	0.02
10		0.06	0.01
11			0.01

Light shaded area indicates over prediction of arrest.

Dark shaded area indicates underprediction of arrest.

No shading indicates prediction accuracy about the same for low and high risk groups.



The first analysis looked at the accuracy of prediction for various categories for low risk when high risk was held constant. The second analysis was to divide the offenders into two groups of risk to examine the accuracy of prediction for various cut off points of high risk when low risk was held constant. Table 9 shows the accuracy of prediction for the various cut-off points for high risk when the low risk group was defined with four different ranges of risk scores: 0-2, 0-3, 0-4, and 0-5. In our previous analysis of two categories, we found that for all of these categories, over 70% of the offenders predicted to be low risk actually were low risk [see Table 4]. Thus, these categories seemed reasonable to set as low risk categories for this analysis.

Table 9. Comparison of Different Cutoff Points and Error Rates for Original and New Burgess Risk Scales, Among Development Sample Offenders. Outcome Defined as Arrest vs. No Arrest (N = 18,261).

	% Forecast high risk	% Forecast medium risk	% Forecast low risk	Total percent of forecasts that were accurate ¹	Of those forecast to be high risk, % that actually were	Of those forecast to be low risk, % that actually were	Of those actually High Risk, % forecasted accurately	Of those actually Low Risk, % forecasted accurately	Overall false positive rate ²	Overall false negative rate ³	Overall false positives /false negatives	False positive/ false negatives , High risk ⁴	False positive/ false negatives , Low risk ⁵
Variations on high risk cutoff value (Low risk: 0-2)													
3 groups: (L: 0-2 H: 3-14)	97.9%	0.0%	2.1%	54.5%	53.8%	88.4%	99.5%	4.0%	46.2%	11.6%	183.47	183.47	0.01
3 groups: (L: 0-2 H: 4-14)	94.2%	3.7%	2.1%	55.8%	55.0%	88.4%	99.5%	4.2%	45.0%	11.6%	171.87	171.87	0.01
3 groups: (L: 0-2 H: 5-14)	88.0%	9.8%	2.1%	57.7%	57.0%	88.4%	99.5%	4.7%	43.0%	11.6%	153.62	153.62	0.01
3 groups: (L: 0-2 H: 6-14)	79.6%	18.3%	2.1%	60.3%	59.5%	88.4%	99.5%	5.5%	40.5%	11.6%	130.73	130.73	0.01
3 groups: (L: 0-2 H: 7-14)	70.3%	27.6%	2.1%	62.7%	62.0%	88.4%	99.4%	6.6%	38.0%	11.6%	108.42	108.42	0.01
3 groups: (L: 0-2 H: 8-14)	58.7%	39.2%	2.1%	65.7%	64.8%	88.4%	99.4%	8.3%	35.2%	11.6%	83.67	83.67	0.01
3 groups: (L: 0-2 H: 9-14)	43.2%	54.7%	2.1%	69.2%	68.2%	88.4%	99.2%	12.0%	31.8%	11.6%	55.62	55.62	0.02
3 groups: (L: 0-2 H: 10-14)	26.2%	71.7%	2.1%	73.7%	72.5%	88.4%	98.7%	20.6%	27.5%	11.6%	29.22	29.22	0.03
3 groups: (L: 0-2 H: 11-14)	11.3%	86.6%	2.1%	77.7%	75.7%	88.4%	97.2%	40.6%	24.3%	11.6%	11.13	11.13	0.09
3 groups: (L: 0-2 H: 12-14)	2.5%	95.3%	2.1%	83.8%	80.0%	88.4%	89.2%	78.6%	20.0%	11.6%	2.07	2.07	0.48
Variations on high risk cutoff value (Low risk: 0-3)													
3 groups: (L: 0-3 H: 4-14)	94.2%	0.0%	5.8%	56.6%	55.0%	81.2%	97.9%	10.0%	45.0%	18.8%	38.67	38.67	0.03
3 groups: (L: 0-3 H: 5-14)	88.0%	6.1%	5.8%	58.5%	57.0%	81.2%	97.9%	11.1%	43.0%	18.8%	34.57	34.57	0.03
3 groups: (L: 0-3 H: 6-14)	79.6%	14.6%	5.8%	61.0%	59.5%	81.2%	97.7%	12.8%	40.5%	18.8%	29.42	29.42	0.03
3 groups: (L: 0-3 H: 7-14)	70.3%	23.9%	5.8%	63.4%	62.0%	81.2%	97.5%	15.0%	38.0%	18.8%	24.40	24.40	0.04
3 groups: (L: 0-3 H: 8-14)	58.7%	35.5%	5.8%	66.3%	64.8%	81.2%	97.2%	18.7%	35.2%	18.8%	18.83	18.83	0.05
3 groups: (L: 0-3 H: 9-14)	43.2%	51.0%	5.8%	69.8%	68.2%	81.2%	96.4%	25.7%	31.8%	18.8%	12.52	12.52	0.08
3 groups: (L: 0-3 H: 10-14)	26.2%	68.0%	5.8%	74.1%	72.5%	81.2%	94.5%	39.7%	27.5%	18.8%	6.58	6.58	0.15
3 groups: (L: 0-3 H: 11-14)	11.3%	82.9%	5.8%	77.6%	75.7%	81.2%	88.6%	63.3%	24.3%	18.8%	2.51	2.51	0.40
3 groups: (L: 0-3 H: 12-14)	2.5%	91.6%	5.8%	80.8%	80.0%	81.2%	65.0%	90.3%	20.0%	18.8%	0.47	0.47	2.15
Variations on high risk cutoff value (Low risk: 0-4)													
3 groups: (L: 0-4 H: 5-14)	88.0%	0.0%	12.0%	59.4%	57.0%	77.1%	94.8%	19.6%	43.0%	22.9%	13.83	13.83	0.07
3 groups: (L: 0-4 H: 6-14)	79.6%	8.5%	12.0%	61.8%	59.5%	77.1%	94.5%	22.3%	40.5%	22.9%	11.77	11.77	0.08
3 groups: (L: 0-4 H: 7-14)	70.3%	17.8%	12.0%	64.2%	62.0%	77.1%	94.1%	25.7%	38.0%	22.9%	9.76	9.76	0.10
3 groups: (L: 0-4 H: 8-14)	58.7%	29.4%	12.0%	66.9%	64.8%	77.1%	93.3%	30.9%	35.2%	22.9%	7.53	7.53	0.13
3 groups: (L: 0-4 H: 9-14)	43.2%	44.9%	12.0%	70.2%	68.2%	77.1%	91.5%	40.2%	31.8%	22.9%	5.01	5.01	0.20
3 groups: (L: 0-4 H: 10-14)	26.2%	61.9%	12.0%	73.9%	72.5%	77.1%	87.4%	56.2%	27.5%	22.9%	2.63	2.63	0.38
3 groups: (L: 0-4 H: 11-14)	11.3%	76.7%	12.0%	76.4%	75.7%	77.1%	75.8%	77.1%	24.3%	22.9%	1.00	1.00	1.00
3 groups: (L: 0-4 H: 12-14)	2.5%	85.5%	12.0%	77.6%	80.0%	77.1%	42.7%	94.8%	20.0%	22.9%	0.19	0.19	5.38
Variations on high risk cutoff value (Low risk: 0-5)													
3 groups: (L: 0-5 H: 6-14)	79.6%	0.0%	20.4%	62.2%	59.5%	72.8%	89.5%	31.6%	40.5%	27.2%	5.80	5.80	0.17
3 groups: (L: 0-5 H: 7-14)	70.3%	9.3%	20.4%	64.4%	62.0%	72.8%	88.7%	35.8%	38.0%	27.2%	4.81	4.81	0.21
3 groups: (L: 0-5 H: 8-14)	58.7%	20.9%	20.4%	66.9%	64.8%	72.8%	87.3%	41.9%	35.2%	27.2%	3.71	3.71	0.27
3 groups: (L: 0-5 H: 9-14)	43.2%	36.4%	20.4%	69.7%	68.2%	72.8%	84.1%	52.0%	31.8%	27.2%	2.47	2.47	0.41
3 groups: (L: 0-5 H: 10-14)	26.2%	53.4%	20.4%	72.6%	72.5%	72.8%	77.3%	67.4%	27.5%	27.2%	1.30	1.30	0.77
3 groups: (L: 0-5 H: 11-14)	11.3%	68.3%	20.4%	73.8%	75.7%	72.8%	60.6%	84.4%	24.3%	27.2%	0.49	0.49	2.03
3 groups: (L: 0-5 H: 12-14)	2.5%	77.0%	20.4%	73.6%	80.0%	72.8%	26.8%	96.7%	20.0%	27.2%	0.09	0.09	10.91

Note. Low risk means that we predicted no arrest. Medium and high risk mean that we predicted any arrest. Only totals for low and high risk offenders were used to predict forecast accuracies. Shaded areas indicate where the error rates are about the same for high risk and low risk categories.

¹ This figure calculated by taking: (low risk offenders who were not arrested + high risk offenders who were arrested)/(total high and low risk offenders)

² False positives are high risk offenders who were not arrested. This figure was calculated by taking: (high risk offenders who were not arrested / total high risk offenders)

³ False negative are low risk offenders who were arrested. This figure was calculated by taking: (low risk offenders who were arrested / total low risk offenders)

⁴ This figure was calculated by taking: (high risk offenders who were not arrested / low risk offenders who were arrested).

⁵ This figure is the inverse of the false positive/false negatives for high risk offenders. It was calculated by taking: (low risk offenders who were arrested / high risk offenders who were not arrested).



Table 10 compares the false positive/false negative ratio for these three risk groups. The non-shaded area shows where there is basically a one to one tradeoff in prediction. For every offender falsely predicted to have an arrest, one offender was falsely predicted to not have an arrest. The light grey shaded area shows the error is more in the direction of over predicting arrest [ratio above 1] and the dark shaded area shows where the error is more likely to be in the direction of under predicting arrest [ratio below 1]. This pattern indicates that to obtain a one to one trade off in the error rate, the high risk category needs to have a different cutoff point depending upon how the low risk category is defined.

Table 10. False positive and false negative rates for Low risk group categories.

Cut off for high risk	Low risk group categories			
	0-2	0-3	0-4	0-5
3	183.47			
4	171.87	38.67		
5	153.62	34.57	13.83	
6	130.73	29.42	11.77	5.80
7	108.42	24.40	9.76	4.81
8	83.67	18.83	7.53	3.71
9	55.62	12.52	5.01	2.47
10	29.22	6.58	2.63	1.30
11	11.13	2.51	1.00	0.49
12	2.07	0.47	0.19	0.09

Light shaded area indicates over prediction of arrest.

Dark shaded area indicates under prediction of arrest.

No shading indicates prediction accuracy about the same for low and high risk groups.

Number of offenders in risk groups. In addition to the prediction accuracy of the various risk group categories, it is also important to know the number of offenders in the categories. This provides an idea of how many offenders may potentially be diverted from incarceration or receive a shorter sentence [low risk] and how many may potentially become incarcerated or receive a longer sentence [high risk]. For example, in our development sample, only three offenders had the highest score of 14 and nine offenders had the lowest risk score of 0. However, there were over 1,000 offenders in each of the categories ranging from 4 through 11. Thus, the risk classification groups that have the greatest accuracy probably have very little impact, due to having so few offenders in the low and/or high risk groups.

Table 11 shows the number of offenders that would be in the low, medium, and high risk categories for eleven different options. The analysis discussed above allowed for a detailed examination of the high risk group and the low risk group to determine where appropriate cut-off points may lie for those two groups. The medium risk group is then determined by default. The options shown in Table 11 include only those risk categories in which the analyses indicated low error rates of either over prediction or under prediction of arrest.

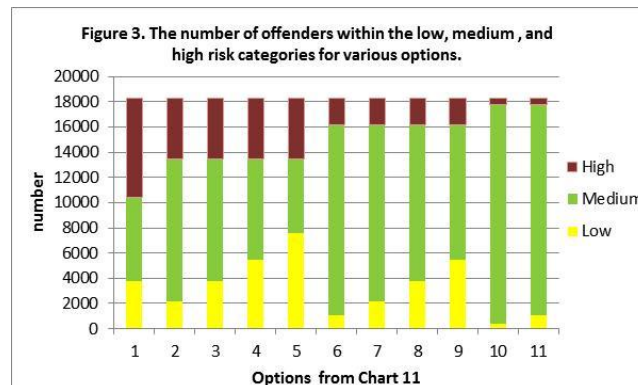


Table 11. The number of offenders in risk classification categories that have low error rates.

Possible risk category options	Risk Scores			Number in risk categories			Percent in risk categories			Number Arrested			Percent Arrested		
	Low	Medium	High	Low	Medium	High	Low	Medium	High	Low	Medium	High	Low	Medium	High
1	0-5	6-8	9-14	3730	6,649	7,882	20.4%	36.4%	43.2%	1015	3269	5379	27%	49%	68%
2	0-4	5-9	10-14	2185	11,295	4,781	12.0%	61.9%	26.2%	500	5697	3466	23%	50%	72%
3	0-5	6-9	10-14	3730	9,750	4,781	20.4%	53.4%	26.2%	1015	5182	3466	27%	53%	72%
4	0-6	7-9	10-14	5432	8,048	4,781	29.7%	44.1%	26.2%	1713	4484	3466	32%	56%	72%
5	0-7	8-9	10-14	7550	5,930	4,781	41.3%	32.5%	26.2%	2717	3480	3466	36%	59%	72%
6	0-3	4-10	11-14	1064	15,134	2,063	5.8%	82.9%	11.3%	200	7901	1562	19%	52%	76%
7	0-4	5-10	11-14	2185	14,013	2,063	12.0%	76.7%	11.3%	500	7601	1562	23%	54%	76%
8	0-5	6-10	11-14	3730	12,468	2,063	20.4%	68.3%	11.3%	1015	7086	1562	27%	57%	76%
9	0-6	7-10	11-14	5,432	10,766	2,063	29.7%	59.0%	11.3%	1713	6388	1562	32%	59%	76%
10	0-2	3-11	12-14	387	17,409	465	2.1%	95.3%	2.5%	45	9246	372	12%	53%	80%
11	0-3	4-11	12-14	1064	16,732	465	5.8%	91.6%	2.5%	200	9091	372	19%	54%	80%

Note: The shaded areas are the options that had the lowest error rate ratios.

Figure 3 shows the number of offenders with the three risk categories for the eleven different classifications from Table 11. This figure shows that there is a great deal of variation in the number of offenders in each risk categories depending how each of the categories is defined. For example, over 90% of the offenders fall within the medium range in options 10 and 11 where the high and low risk categories are narrowly defined. On the other hand, option 1 has the plurality of offenders in the high risk group since that option has high risk broadly defined to include six of the 14 risk scores. Most of the other options do not have these extremes. Rather, the majority of offenders fall within the medium risk category [ranging from 53% to 83%] with a more balanced split between the high and low risk categories. They do, however, still provide for some latitude in the decisions to be made about whether it is better to potentially err on the side of over prediction of arrest [more offenders in high risk category by lowering the cutoff score for high risk] or under prediction of arrest [more offenders in low risk group by raising the cutoff score for low risk].





Policy Decisions.

In the final determination of the risk classification groups there are two policy issues to be considered:

1. *Potential Impact of risk categories on sentencing practices.* It is important to know how many people would potentially be in the risk categories and to know the impact of the risk assessment on sentencing practices. For example, how many offenders would be projected to receive a harsher sentence [resulting in more incarceration] compared to how many people would be projected to receive a more lenient sentence [resulting in less incarceration]. What would be the overall net impact of the various risk classification categories?
2. *False positive/false negative error tradeoffs.* Is it better to err on the side of over predicting arrest [which potentially could result in correctional overcrowding] or under predicting arrest [which potentially could result in more crime]. How much better? That is, how many false positives equal one false negative [or vice versa]?

While we have completed the prediction and error rate analyses, we will next be examining the potential impact associated with the various risk score classification categories. At that point, the Commission will have the necessary information to address the above policy issues.



Appendix Table 1A. Logistic regression predicting 3-year recidivism among development sample offenders (N = 18,239).¹

	Solitary offenders	Multiple offense offenders	Full development sample	Full development sample
	Odds ratio	Odds ratio	Odds ratio	Odds ratio
OGS < 4²	1.165**	1.246***	1.198***	1.203***
1 prior arrest³	1.528***	1.450***	1.487***	1.486***
2-4 prior arrests³	2.394***	1.945***	2.146***	2.145***
5-12 prior arrests³	3.467***	2.901***	3.155***	3.155***
13+ prior arrests³	6.162***	4.709***	5.372***	5.365***
Prior property arrest	1.196**	1.182*	1.191***	1.192***
Prior drug arrest	1.209***	1.143*	1.178***	1.179***
Personal offense ⁴	1.033	0.907	0.968	0.967
Property offense ⁴	1.227***	1.087	1.158***	1.156***
Firearms/weapons offense ⁴	0.909	0.938	0.927	0.927
Other offense ⁴	1.126	0.687***	0.927	0.93
24-29 years⁵	0.647***	0.579***	0.611***	0.611***
30-49 years⁵	0.414***	0.380***	0.397***	0.397***
50+ years⁵	0.173***	0.169***	0.171***	0.172***
Female	0.748***	0.844*	0.794***	0.794***
Black⁶	1.379***	1.297***	1.340***	1.341***
Hispanic ⁶	1.242*	1.12	1.183**	1.183**
Other race ⁶	1.042	0.795	0.921	0.922
Other urban county⁷	0.828***	0.867**	0.847***	0.846***
Rural county⁷	0.666***	0.752***	0.705***	0.705***
Multiple charges				1.039
N	9442	8645	18087	18087
Pseudo R-squared	0.102	0.084	0.092	0.092
AIC	11774.7	11007.8	22775.4	22776
BIC	11924.9	11156.1	22939.2	22947.6

* p < .05 ** p < .01 *** p < .001

Note. Variables in bold are significant at p < .05.

¹ Does not include offenders whose first failure was a technical violation, who did not have a 3-year follow-up, and who were missing gender or race.

² Reference is OGS > 3

³ Reference is no prior arrests

⁴ Reference is drug offense

⁵ Reference is less than 24 years

⁶ Reference is white

⁷ Reference is Philadelphia/Allegheny County