

Sonoma County Static Risk and Needs Assessment Tool Validation Study Results

March 26, 2025

Dr. Hollie MacDonald, Senior Researcher, DNA Global

Dr. Sonia Jain, Founder and Principal, DNA Global

www.datainaction.org

6 Program Objectives

Sonoma County CCP Public Safety Realignment Implementation Plan 2023-2024

1. **Reduce recidivism** to enhance public safety.
2. Promote **evidence-based programming** and upstream investments in health, education, and human services to decrease the need for and costs of enforcement, prosecution, and incarceration.
3. **Fund programs** that align with tenets of [Sonoma County's Criminal Justice Master Plan](#).
4. Minimize use of jail beds through use of **detention alternatives** in a manner that is consistent with public safety and maintains integrity of the criminal justice system.
5. Provide programming for in-custody and out-of-custody individuals and **use validated risk assessments** to inform programming decisions and ensure continuity.
6. **Operate a day reporting center** to serve as central point of evidence-based programming to help justice-involved individuals reintegrate into the community.

Current Evaluation Priorities by the CCP

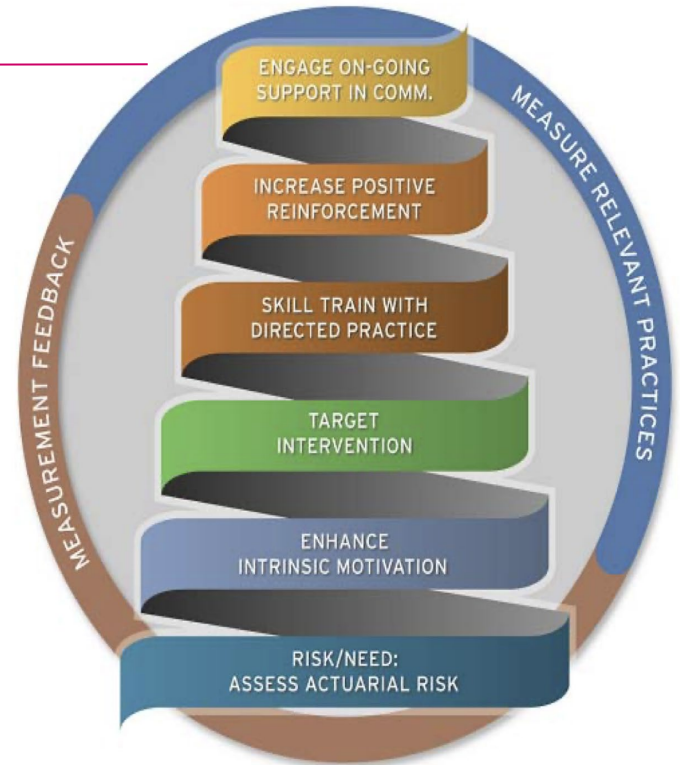
1. **Local Validation of the Static Risk Assessment and Offender Needs Assessment used for classification and case planning for people under Probation supervision.**
2. Race and gender disparities analysis, comparing proportions of race and gender groups at key justice system decision points. *(next up)*
3. Examine effectiveness of electronic monitoring in supporting arrest-free behavior and court appearance.
4. Evaluation and planning to identify and address disparities related to race, gender, and mental health status in program engagement and outcomes.
5. Process and outcome evaluation of new substance use disorder services at Probation Day Reporting Center once the program has been in operation long enough to be evaluated.
6. Implementation assessment of Probation's behavior response sy

Why Validate?

Ensuring Evidence Based Practice



1. Reliability
 - Norm/Repeatability/Consistency
 - Inter-rater and Test-retest
 - Prerequisite to Validity
2. Validity
 - Accuracy of Tool
 - Knowing the Strongest Predictors
 - = Cost Effective Practices



Evaluation Questions 1 and 2: **Fidelity and Reliability**

- 1. Is the Sonoma County SRNA Tool being used with fidelity?**
 - a. Regarding the SRA and ONA, how much confidence should we have in the assessments' performance?
 - b. How well does the SRNA tool risk classification align with supervision assignment?
 - c. What does each risk classification actually mean regarding the SRA?

- 2. Are staff accurately and consistently completing the ONA assessments?**
 - a. Are the Sonoma County SRNA scoring practices reliable?

Evaluation Question 3: Predictive Validity

3. How well does the Sonoma County SRNA Tool assess the risk of recidivism overall?
 - a. Does predictive validity **vary by risk categories versus continuous risk score**?
 - b. Does predictive validity **vary by time to recidivist event** when accounting for any jail time during the supervision period?
 - c. What **need and strength areas** are strongly associated with re-offense and its avoidance as outlined in the ONA?
 - d. How often are **overrides** happening? Does this impact outcomes?
 - e. Should we reinforce or **adjust the SRA risk level cut points** to classify people most effectively, such that intensity of supervision aligns with risk to reoffend?

Evaluation Question 4: Subgroups

4. How well does the SRNA Tool perform for specific subgroups?

- a. Are the SRA and ONA assessments' effective with various race, gender and age groups? Are adjustments needed to make it more effective?
- a. How well does the SRNA Tool perform in the presence of a mental health need?
- a. Does predictive validity vary by type of offense (such as domestic violence, DUI, sex offender)?
 - Confirm that the SRNA Tool does not accurately capture some unique offense types such as DV, DUI, and sex offenders.



Data and Methodology



Data and Sample

Excellent and comprehensive, Sonoma County Probation Records data, 2012-2024. Merged risk-recidivism data, Matched clients across files using unique identifier.

Final Sample= **19,442, high quality**

Specifically:

1. **Static Risk Assessment (SRA) Instances**
2. **Offender Needs Assessment (ONA) Instances**
3. **Reoffenses Data (Rearrest only, Rearrest with Conviction, Type of Rearrest)**
4. **Jail Data (Time Spent in Jail)**
5. **Override Data (Presence of Overrides, Up, Down or Hold/Tx and Reasons)**
6. **Advanced Statistical Models:**
 - SRA: Risk Classification and Category Scores
 - ONA: Factor Scores, Risk and Protective Scores, individual item scores, constellations of item scores
 - effectiveness for subgroups, accuracy and consistency of assessments, jail time during observation period, override presence, predictive and protective factors (ONA).
 - **Highly Technical, Exploratory and Confirmatory analysis. Re-ran models to incorporate multiple IVs and DVs.**

Dependent Variable: **Recidivism within 1 year, 3 years, and ever**

- **Recidivism within one year** - Primary outcome, presence of conviction date in reoffense data within 365 days of initial assessment
- **Recidivism within three years** - Presence of conviction date within 1,095 days.
- **General Recidivism (ever)** - Presence of conviction date in reoffense data anytime after SRA assessment.
- **Time to Recidivism** - Days Spent in the community between SRNA Assessment date and the first rearrest date. Also considered **jail time**.
- **Rearrest** - Also considered as outcome to examine effectiveness of risk levels in predicting arrests that did NOT end in conviction

Independent Variables: Risk and Needs Factors



- **Risk Scores:** Felony Risk Score, Property Violent Risk Score, Violent Risk Score
- **Five Risk Levels:** Low, Moderate, High Drug, High Property, High Violent
- **Three Risk Levels:** Low, Moderate, High Combined
- **Overrides:** Any Override, Override Up, Override Down, Race, Gender, Age Group,
- **Initial Offense Type:** Felony Drug Offense, Felony Property Offense, Felony Violent Offense, Misdemeanor, Others
- **Rearrest Type:** Arrest only or conviction
- **Mental Health Risk:** Mental Health Status and Suicide Risk

Models predicted validity to 5 outcomes: **overall recidivism** (conviction), recidivism within one year (conviction), recidivism within three years (conviction), **rearrest** (no conviction), and **time to any rearrest** (either arrest only or conviction), **rearrest offense**, and **type of rearrest** (Felony drug arrest, Felony property crime arrest and Felony Violent rearrest, etc).

Statistical Data Analyses

Fidelity Analysis:

- Multiple Regression Models
- Cross-tabs and ANOVAs
- Logistic Regression Models and Nagelkerke R^2 values, Adjusted Odds Ratios

Reliability:

- Two-way mixed-effects model both SRA and ONA
- Cronbach's Alpha for SRA
- Intraclass Correlation Coefficient (ICC) analyses for SRA and ONA
 - The single measures ICC assesses the reliability of a single coder's ratings and the average measures ICC estimates reliability when multiple coders' scores are averaged.
 - The SRA is 100% automated, so reliability is examining the consistency of scoring rather than rater decision-making.

Predictive Validity:

- Logistic regression models
- Survival Analyses
- Receiver Operating Characteristics (ROC) & Area Under the Curve (AUC) Statistics, Youden J Index, Gini Index, K-S Stats
- Kaplan Meier Survival Tests, and Cox Proportional Hazards Models.

Overrides Analysis:

- Area Under the Curve (AUC) Statistics and Kaplan Meier Survival Tests.

Subgroup Analysis:

- Logistic regression models
- Survival Analyses
- Receiver Operating Characteristics (ROC)

Evaluation/Validation Study Aims and Goals:



Validation Study Aims:

- To validate the Static Risk Assessment (SRA) and Offender Needs Assessment (ONA) tools overall and by subgroups, used by Sonoma County Probation.

Primary Goals:

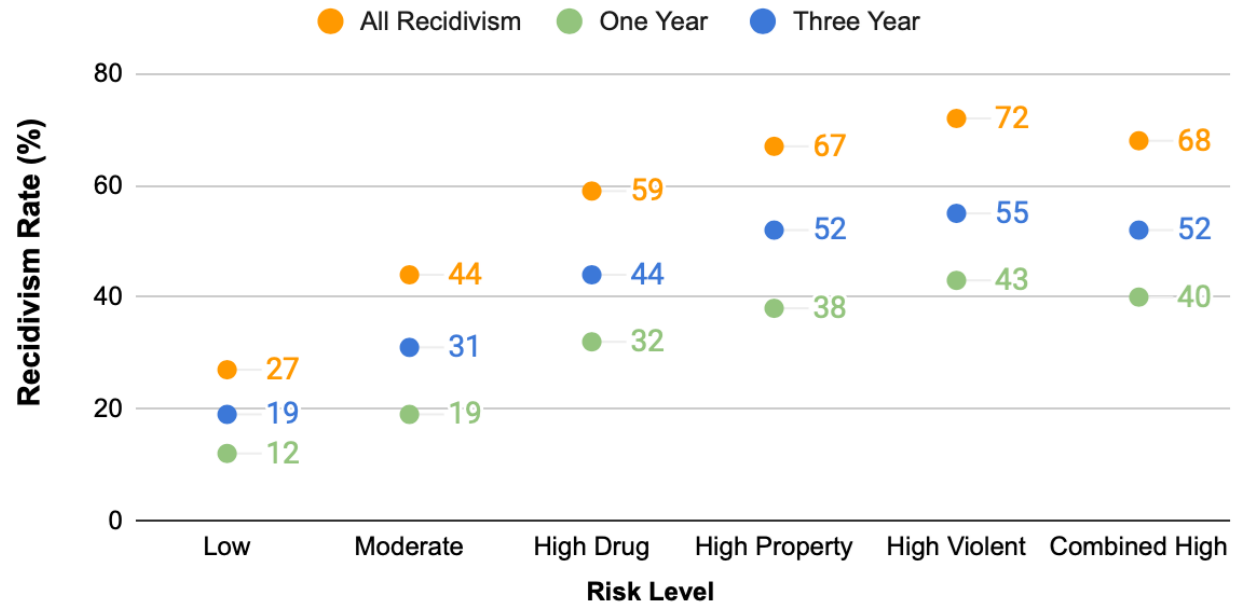
- To assess the tools' fidelity, predictive performance, and fairness across demographic groups while examining their alignment with evidence-based supervision practices.

"Q1: Is the SRNA Tool Used with Fidelity?"

Yes.

- SRNA felony, property violent, and violent scores **significantly predict risk level**, as designed ($p < .001$).
- Risk level classification shows **moderate to strong predictive validity** for general recidivism, with higher risk levels correlating with increased recidivism rates (e.g., 27% for low-risk vs. 72% for high-risk).

Figure 3. Recidivism Rate of SCPD Clients by Risk Level (2012-2024)



Data Source: Combined SRNA and Reoffense Data (n=19442)

Q1 Continued

b. How well does SRNA tool risk classification align with supervision assignment?

- **Risk classification aligns well with supervision.** Significant differences in recidivism across risk levels, **higher risk predicts more reoffending** and post-hoc tests confirm these differences, **supporting its use for supervision decisions.**

c. What does each risk classification actually mean regarding the SRA?

- The tool **effectively predicts recidivism risk**, esp. for drug and property offenses. **Weaker for Misdemeanors, Violent Felonies, and Arrests without Conviction**

Low Risk (Level 1)	1 year 12%	3 year 19%	All 27%	These individuals may require minimal supervision and could benefit from low-intensity interventions or monitoring.
Moderate Risk (2)	19%	31%	44%	May require moderate supervision and targeted interventions to reduce recidivism risk.
High Drug Risk (3)	32%	44%	59%	May requires substance use interventions, drug courts, and intensive supervision.
High Property Risk (4)	38%	52%	67%	May benefit from restitution programs, behavioral interventions and intensive supervision.
High Violent Risk (5)	43%	55%	72%	Requires strict supervision, behavioral therapy, and structured reentry programs.

Q2. Reliability Analysis

2a. Are staff accurately and consistently completing the ONA assessments?

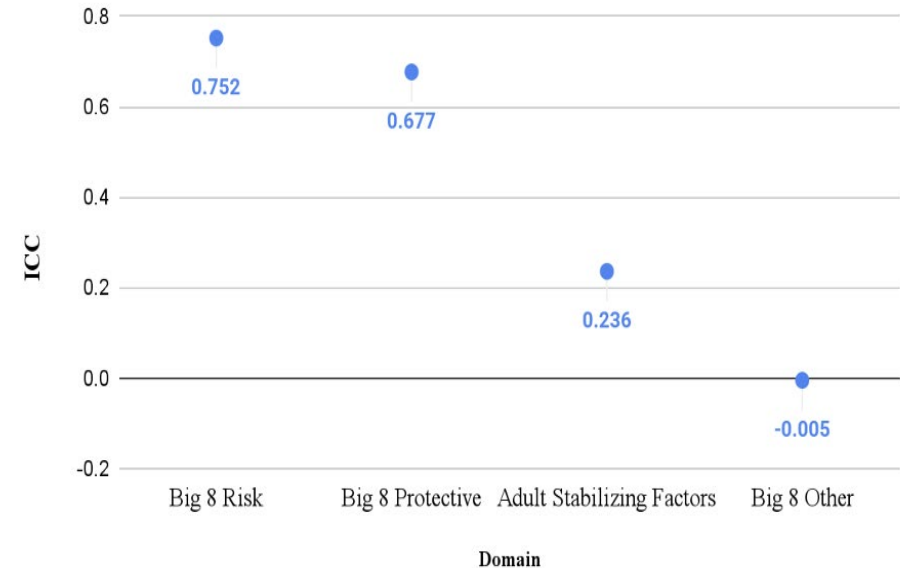
ONA assessment accuracy varies by domain:

- **Big 8 Risk: Acceptable** consistency and **moderate** reliability.
- **Big 8 Protective:** Moderate consistency but lower reliability.
- **Adult Stabilizing & Big 8 Other:** Poor reliability but this is due to scoring mechanism not actual scorers (not used for supervision setting, just for resource decision making).

2b. Is the Sonoma County SRNA reliable across raters (inter-rater reliability) and over time (test-retest reliability)?

- **High inter-rater reliability**
- Variation in scores reflects true differences in individuals, not rater inconsistencies.

Figure 4. Reliability of ONA Domain Scores



Q3. SRNA and Risk of Recidivism

a. How well does the Sonoma County SRNA Tool assess risk of recidivism overall?

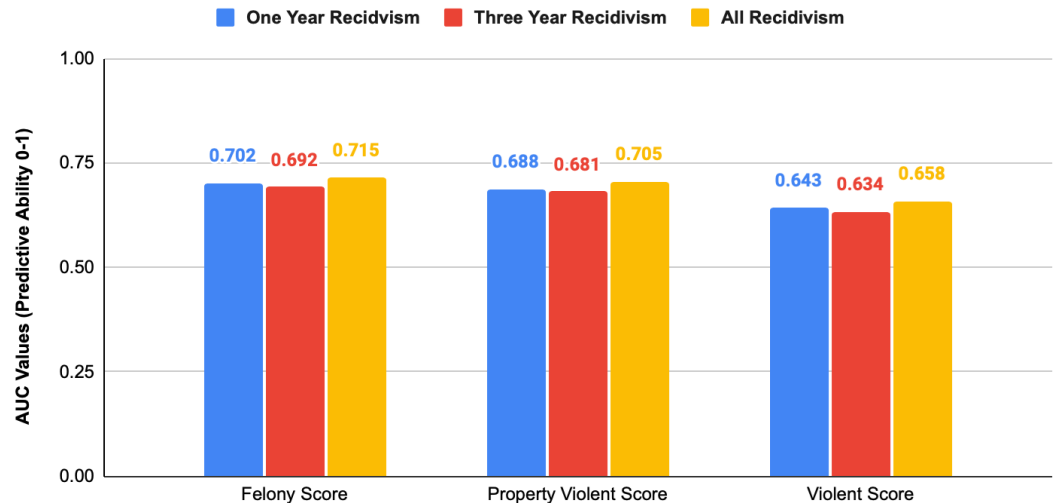
- **Moderate predictive validity.** **Felony Score** is **strongest** predictor of recidivism (AUC = 0.715), but no score is highly accurate alone. Predictive power could improve by incorporating additional risk factors and validated tools.

b. Does predictive validity vary by risk categories versus continuous score?

Figure 5. Predictive Ability of Risk Scores on Recidivism Rate of SCPD Clients (2012-2024)

Ranked Predictive Ability of Risk Categories vs. Risk Scores

1. **Felony Score = .702, .692, .715**
2. **Prop Violent Score = .688, .681, .705**
3. **Five Risk Levels = .683, .683, .694**
4. **Three Risk Levels = .676, .667, .688**
5. **Violent Score = .643, .634, .658**

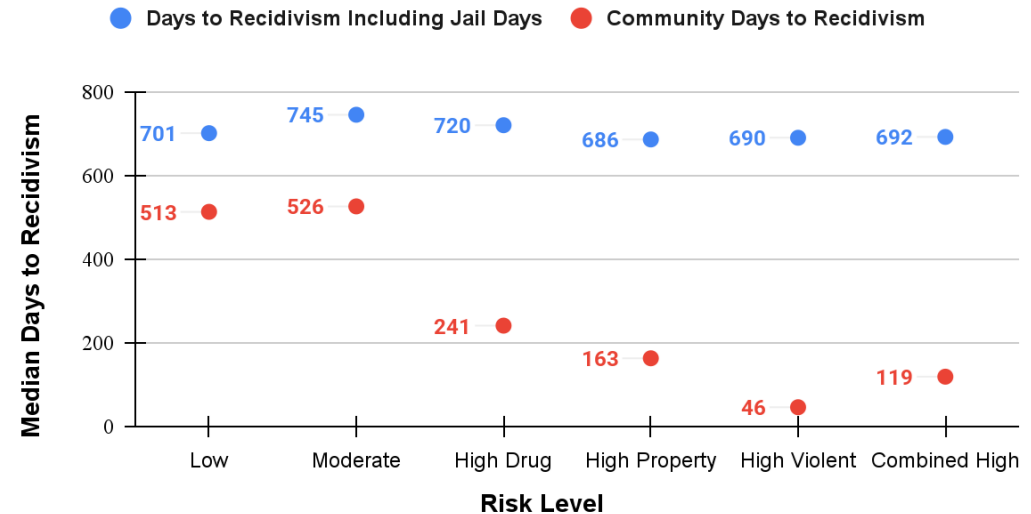


Q3. Does **predictive validity** vary by time to recidivist event, when accounting for any jail time during the supervision period?

Yes, predictive **validity** changes when accounting for jail time.

- Including jail time extends time to recidivism across all risk levels, resulting in more uniform survival times.
- This suggests that while jail time may temporarily delay recidivism, **it does not significantly alter the underlying risk factors or prevent recidivism.**

Figure 6. Median Days to Recidivism for SCPD Clients (Both Models)



Data Source: Combined SRNA, ONA, and Reoffense Data (n=19442)

Q3. Needs and Strengths - Re-offense analysis

C. What need and strength areas are strongly associated with re-offense and its avoidance as outlined in the ONA?

Table 6. Factors That Predict or Protect Against Recidivism Alone

Factor and Factor Domain	B	β	Significance	Impact
F1: Antisocial Personality - Big 8 Risk	0.11	1.011	<.001	Predictive
F2: Antisocial Behavior - Big 8 Risk	.005	1.005	<.001	Predictive
F3: Criminal Thinking - Big 8 Risk	.013	1.013	<.001	Predictive
F4: Criminal Associates - Big 8 Risk	.011	1.011	<.001	Predictive
F5: Family - Big 8 Risk	.022	1.022	<.001	Predictive
F6: Employment/School - Big 8 Risk	.015	1.015	<.001	Predictive
F8: Substance Abuse - Big 8 Risk	.014	1.014	<.001	Predictive
F9: Antisocial Personality - Big 8 Protect	-0.008	.992	<.001	Protective
F10: Antisocial Behavior - Big 8 Protect	-0.005	.995	<.001	Protective
F11: Criminal Thinking - Big 8 Protect	-0.008	.992	<.001	Protective
F12: Criminal Associates - Big 8 Protect	-0.010	.990	<.001	Protective
F13: Family - Big 8 Protect	-0.009	.991	<.001	Protective
F14: Employment/School - Big 8 Protect	-0.015	.985	<.001	Protective
F16: Substance Abuse - Big 8 Protect	-0.008	.992	<.001	Protective
F18: Antisocial Behavior - Big 8 Other	.003	1.003	<.001	Predictive
F21: Family - Big 8 Other	-0.003	.997	.331	Not Sig
F22: Employment/School - Big 8 Other	.005	1.005	.007	Predictive
F24: Substance Abuse - Big 8 Other	.035	1.035	<.001	Predictive
F50: Mental Health Issues - Adult Stability	.006	1.006	<.001	Predictive
F52: Homeless/Runaway - Adult Stability	.014	1.014	<.001	Predictive
F53: Current Abuse/Neglect- Adult Stability	.009	1.009	<.001	Predictive
F54: Medical Issues - Adult Stability	.006	1.006	<.001	Predictive

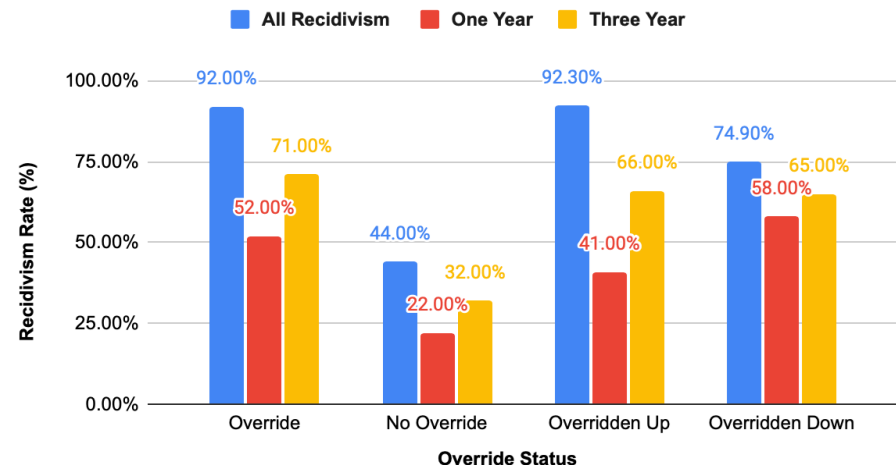
Q3. Override Analysis

d. How often are overrides happening? Do overrides impact outcomes?

- Overrides occur in 8% of cases: 8.5% shift to higher risk, 3.2% to lower, rest lateral or holds.
- **Overrides impact outcomes significantly: Overall, overridden cases have 92% recidivism rate, vs. 44% non-overridden cases.**

- **Upward Overrides** (increased risk classification) had lower short-term recidivism but higher long-term recidivism (66% over three years).
- **Downward Overrides** (reduced risk classification) had higher short-term recidivism (58% within one year), suggesting an underestimation of risk.
- Survival analysis confirms that downward overrides significantly shorten the time to re-offending, reinforcing concerns about risk underestimation.

Figure 7. Override Status on Recidivism of SCPD Clients (2012-2024)

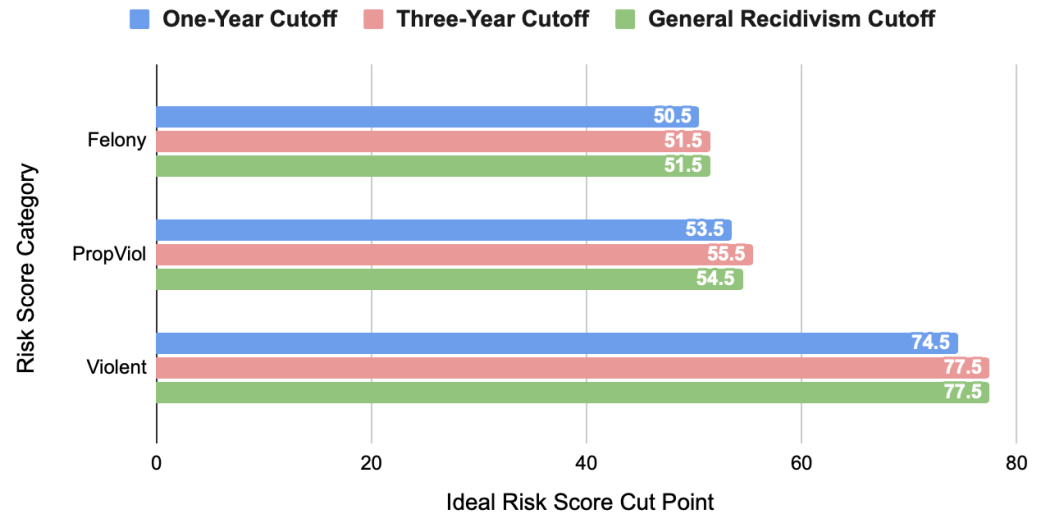


Data Source: Combined SRNA and Reoffense Data (n=19442)

Q3. Adjusting cut-off points (cont)

- e. **Should we reinforce or adjust the SRA risk level cut points to classify people most effectively, such that intensity of supervision aligns with risk to reoffend?**
 - Adjusting SRA cutoff scores could improve accuracy for both predicting general recidivism, one-year, three-year and offense specific risk.
 - The most precise cut points also differ for gender and race (see Q4 results).

Figure 10. Most Precise Risk Score Cut Off Points for SCPD Clients



Data Source: Combined SRNA and Reoffense Data (n=19442)

Q3. Specific Offense Cut Points (General Recidivism)

Table 11. Specific Cut Points for Specific Offenses are noted in the table below.

Classification Level	Current Cut Point	Ideal Cut Point
High Violent (Violent Offense)	96	82.5
High Property (Property Offense)	66	63.5
High Drug (Drug Offense)	66	61.5
Moderate Risk (Property/Violent Offense)	50	58.5
Moderate Risk (Felony Offense)	50	50.5

- **Ideal cut points differ by goal of the prediction**, whether its specific to offense type (as the scoring mechanism is designed) or if it is any recidivism type

Q4. Race/Ethnicity Analysis

4a. Are the SRA and ONA assessments' effective with various race, gender and age groups? Are adjustments needed to make it more effective?

- Most accurate for White and "Other" racial groups (Asian, Pacific Islander, American Indian)
- Most precise cutoffs differ between racial groups - Blacks recidivate with higher scores than Whites

Table 13. Predictive Validity by Subgroups (Race)

	<i>White</i>	<i>Ideal Cut Off</i>	<i>Black</i>	<i>Ideal Cut Off</i>	<i>Hispanic</i>	<i>Ideal Cut Off</i>	<i>Other</i>
General Recidivism							
Felony Score	AUC = .727	51.5	AUC = .688	64.5	AUC = .696	51.5	AUC = .758
Property Violent Score	AUC = .715	54.5	AUC = .673	62.4	AUC = .696	55.5	AUC = .753
Violent Score	AUC = .669	76.5	AUC = .665	87.5	AUC = .642	77.5	AUC = .691
One Year Recidivism							
Felony Score	AUC = .708	54.5	AUC = .679	64.5	AUC = .697	51.5	AUC = .725
Property Violent Score	AUC = .691	58.4	AUC = .663	67.5	AUC = .689	55.5	AUC = .718
Violent Score	AUC = .647	78.5	AUC = .633	86.5	AUC = .643	84.5	AUC = .638

Source: Validation of the Sonoma County Adult Probation Risk and Needs Assessment Tool, February 2025.

Q4. Gender

4a. Are the SRA and ONA assessments' effective with various race, gender and age groups? Are adjustments needed to make it more effective?

- Performs better for men and ideal cut points differ (somewhat drastically) for men and women.

Table 12. Predictive Validity by Subgroups (Gender)

			Ideal Risk	Score Cut Offs
	<i>Female</i>	<i>Male</i>	<i>Female</i>	<i>Male</i>
General Recidivism				
Felony Score	AUC = .696	AUC = .716	47.5	51.5
Property Violent Score	AUC = .672	AUC = .710	49.5	54.4
Violent Score	AUC = .633	AUC = .681	44.5	77.5
One Year Recidivism				
Felony Score	AUC = .699	AUC = .702	47.5	59.6
Property Violent Score	AUC = .683	AUC = .688	50.5	61.5
Violent Score	AUC = .629	AUC = .64	45.5	85.5

Data Source: Combined SRNA, ONA, and Reoffense Data (n=19442)

Q4. Age

4a. Are the SRA and ONA assessments' effective with various race, gender and age groups? Are adjustments needed to make it more effective?

- Works better for older adults (likely because they have more history to base risk off of).
- Cut-offs vary: Younger offenders require higher risk scores to be classified as high risk, while older offenders re-offend at lower scores.

Table 14. Predictive Validity by Subgroups, ROC Analyses, and Ideal Cut Offs (Age)

<i>General Recidivism</i>	<i>17-19</i>	<i>20-29</i>	<i>30-39</i>	<i>40-49</i>	<i>50-50</i>	<i>60+</i>
Felony Score	AUC = .602	AUC = .700	AUC = .726	AUC = .717	AUC = .720	AUC = .745
Property Violent Score	AUC = .586	AUC = .683	AUC = .720	AUC = .713	AUC = .714	AUC = .728
Violent Score	AUC = .570	AUC = .636	AUC = .652	AUC = .665	AUC = .673	AUC = .715
Ideal Cut Off Felony	50.5	51.5	51.5	50.5	43.5	35.5
Ideal Cut Off Prop Viol	60.5	60.5	53.5	53.5	39.5	24.5
Ideal Cut Off Violent	84.5	84.5	77.5	74.5	60.5	50.5
<i>One Year Recidivism</i>						
Felony Score	.559	.702	.715	.687	.718	.746
Property Violent Score	.534	.693	.716	.686	.717	.723
Violent Score	.476	.655	.646	.649	.676	.720
Ideal Cut Off Felony	53.5	58.5	59.5	51.5	46.5	38.5
Ideal Cut Off Prop Viol	64.5	60.5	54.4	55.5	44.5	24.5
Ideal Cut Off Violent	57	86.5	77.5	74.5	61.5	50.5

Q4. Performance w/ Mental Health need

b. How do the tools perform when there is a presence of mental health need?

- Mental health (like suicide history) increases recidivism risk ($p < .001$)
- Predictive ability is lower when Mental Health History or Suicide Risk is included (.552 - .574)
- Tool more accurate predictive ability when mental health is not considered in the scoring mechanism.
- **Key Takeaway:** Mental health matters for support plans, not as the main focus of risk assessment.

Q4. Varies by type of offense?

c. Does predictive validity vary by type of offense?

- Some offenses are correlated with future recidivism but prior offense type should not be the only consideration when predicting risk.

Table 16. Effect of Felony Offense Type on Recidivism and Tool Validity (2012-2024)

	General Recidivism		One Year Recidivism	
	β	AUC	β	AUC
Felony Prop	1.422	0.596	1.401	0.611
Felony Drug Offense	1.289	0.534	1.275	0.535
Felony Weapon	1.573	0.533	1.759	0.545
Felony DV	1.316	0.511	1.395	0.514
Felony Robbery	1.289	0.51	1.384	0.514
Homicide Offense	.498	0.497	.629	0.498
Felony Sex Offense	.732	0.493	.949	0.498

Preliminary 10 Recommendations

1. Confirm Cutoff Score Refinement for Enhanced Predictive Accuracy

- **Adjust Risk Level Cut Points:**

- Lower **High Violent Risk** cutoff score: **96** to **82.5** to improve sensitivity and balance specificity
- Reduce **High Property Risk** cutoff **66** to **63.5**, improving sensitivity to **68.9%**, thus better identifying high-risk individuals while minimizing false positives.
- Lower **High Drug Risk** cutoff **66** to **61.5**, enhancing classification reliability as indicated by an improved Youden's Index.
- For **Moderate Risk**, increase the **Property/Violent Felony Score** cutoff to **58.5** and adjust the **Felony Score** to **50.5** to reduce over-classification and false positives.

- **Annually/Biennially conduct a simulation** or follow-up study to confirm risk score changes are working precisely.

Preliminary Recommendation #2

2. Calibrate per Racial and Gender Disparities

- **Racial Equity:**
 - Adjust these scoring cutoffs to ensure more equitable risk classification across all racial groups, at the same cut-point.
 - Regularly check outcomes to ensure fairness.
- **Gender Sensitivity:**
 - Lower cutoff points for female clients (e.g., reducing Violent risk cut off from **77.5 to 44.5** for women).
 - Use supplementary risk assessments for female clients to capture risk factors not well-predicted by the SRNA tool.

~~Preliminary Recommendations #3 and #4~~

3. Continue to Use Specialized Tools for Specific Offense Categories

- The SRNA tool exhibits lower predictive accuracy for sex crimes, domestic violence, and DUIs (performing no better than random chance).
- Keep using separate, specialized tools for these crimes to predict risk better.

4. Enhance the Offender Needs Assessment (ONA) for Better Resource Allocation

- The Big 8 Risk and Big 8 Protective domains demonstrate acceptable reliability.
- Adult Stabilizing and Big 8 Other domains should be used for needs assessment and resource provision only (they already are).
- Integrate protective ONA factors like employment stability and housing into individualized supervision plans to reduce recidivism risk

Preliminary Recommendations #5 and #6

5. Integrate Dynamic Risk Factors into SRNA for Enhanced Predictive Power

- Recent employment changes
- Family stability
- And substance abuse patterns

6. Differentiate Between Short-Term (One-Year) and Long-Term (General) Risk Predictions

- Use a higher risk threshold for one-year recidivism predictions to prioritize intervention:
 - Felony Score One-Year Recidivism Cutoff: 54.5 (AUC = 0.702, Youden's Index = 0.321)
 - Property/Violent Score One-Year Cutoff: 58.5 (AUC = 0.688, Youden's Index = 0.296)
- Use slightly lower thresholds for general recidivism to improve long-term predictions:
 - Felony Score General Recidivism Cutoff: 51.5 (AUC = 0.715, Youden's Index = 0.336)
 - Property/Violent Score General Cutoff: 54.5 (AUC = 0.705, Youden's Index = 0.316)

Preliminary Recommendations #7 and #8

7. Reassess the Use of Overrides in Risk Classifications

- Monitor and standardize override decisions and evaluate how often probation officers override SRA scores while emphasizing documentation of the justifications behind these decisions - Most are due to holds but some had no justification noted.
- Limit downward overrides (lowering risk levels) due to a 58% one-year recidivism rate.

8. Implement Continuous Training and Feedback Loops

- Actual Outcomes and Feedback
 - Use **feedback loops** where probation officers can review their assessments against short and long-term recidivism outcomes to improve future scoring accuracy.

Preliminary Recommendations #9 and #10

9. Tailor Interventions Based on Offender Subgroup Characteristics

- Design targeted intervention programs based on subgroup characteristics, such as age, gender, or offense type. For example:
 - Implement gender-responsive interventions for female offenders.
 - Assure programs focused on addressing substance abuse, homelessness, and mental health issues for high-risk groups, as these factors significantly predict recidivism.

10. Account for the Effect of Jail Time on Recidivism Risk

- Recognize that jail artificially extends survival time and does not reduce actual risk.
- Develop risk-adjusted supervision strategies for individuals reentering the community to reflect the delayed but not diminished recidivism risk.

Feedback? Questions?

www.datainaction.org

Next Stage:

Race and gender disparities analysis.

- Comparing proportions of race and gender groups at key justice system decision points