

**STATUS REPORT: VALIDATING THE NEW MEXICO RISK/NEEDS ASSESSMENT
INSTRUMENT**

Prepared for
Probation and Parole Division,
New Mexico Corrections Department

June, 1998

Prepared by
The Center for Applied Research and Analysis
The Institute for Social Research
University of New Mexico

CENTER FOR APPLIED RESEARCH AND ANALYSIS

INSTITUTE FOR SOCIAL RESEARCH

RESEARCH TEAM STAFF

Data Analyst and writing: Kristine Denman, M.A.

Statistical Consultant and writing: Ray Liedka, Ph.D.

Principal Investigator: Paul Guerin, Ph.D.

Data Collection Coordinator: Mercedes Murphy, B.A.

Research Assistants: Gabe Archibeque, B.A.

Rebecca Griego

Stormi Grefphenreed

Rachel Heisler

Sarah Kurhajetz

Felicita Marquart

Audrey Merriweather

Jeff Mix

Brock Perkins, B.A.

Jeannette Valdez, B.A.

INTRODUCTION

This status report is one in a series of deliverables for our current New Mexico Corrections Department (NMCD), Probation and Parole Division (PPD) evaluation of Community Corrections (CC) Program statewide. For reasons beyond our control we were not able to complete the task of validating the current Risk/Needs Assessment instrument used by the NMCD Probation and Parole Division. This status report provides a preliminary analysis of the data collected through this contract period.

New Mexico Corrections Department Probation and Parole Division contracted with the Institute for Social Research's (ISR) Center for Applied Research and Analysis (CARA) to validate the Risk/Needs Assessment (RNA) instrument that is currently being used by Probation and Parole to determine supervision classifications. The risk/needs instrument used in New Mexico is based on the widely used Wisconsin Risk/Needs Assessment Instrument. This instrument contains several sections including both risk and needs assessment forms which establish the basis of a client's supervision classification, and a form for the supervision plan. Many other states have chosen to use the Wisconsin instrument as a result of a push by the National Institute of Corrections to implement an existing risk assessment instrument (Wright, Clear and Dickson, 1984). The NIC would not fund the development of new instruments (ibid). However, to ensure that the instrument was appropriate, the NIC indicated that the adopted instrument should be validated within a few months of its implementation in a new state. This has not been done in New Mexico to date. With the exception of some minor changes in wording, New Mexico has not modified the Wisconsin instrument for use in this State.

Nationally, risk assessment tools have been used to classify offenders since the 1970s. There are several benefits to using a risk prediction instrument. One is that it helps to identify the offenders who pose the greatest risk to the community, thereby protecting the community. It also helps to ensure the proper level of supervision given to offenders and improves the appropriate use of resocialization programs. Unlike "clinical" decision making (one in which the officer determines the level of supervision based on their own judgement), the actuarial model (use of an objective instrument) allows for objective decisions to be made regarding an appropriate level of supervision. However, special circumstances can be accounted for because the classification decision can be overridden. Finally, it is a better use of scarce resources. Offenders are assigned to an appropriate level of supervision better utilizing available funds.

In order to achieve these goals, however, the assessment tool must be valid. Validation means that the instrument does what it is meant to do: it predicts risk of failure while on probation or parole and/or recidivism, thereby determining the appropriate level of supervision. Although the instrument on which New Mexico's risk/needs assessment has been validated, it may not be valid for New Mexico. Validation is specific to geographical location and time because offender populations change over time and place. That is, risk/needs assessments are specific to the population for which they were designed. Thus, a risk/needs instrument that was validated in one geographical area (Wisconsin) may not be valid in another geographical area (New Mexico). Once an instrument is valid, including the validation of all revisions made to it, it must be revalidated periodically for subsequent revisions.

Several steps must be taken before a validation check can be completed, some of which have already been accomplished. First, a review of the literature regarding the use and validation of RNA instruments must occur. Decisions made in subsequent steps are based at least in part on the findings from the literature review. The second step is to resolve methodological concerns including sampling issues (such as what is an appropriate sample and what is an appropriate sample size), selection of criterion and predictor variables, and statistical techniques. Third, the data must be collected and coded. Finally, the analysis must be completed. The remainder of this report includes a discussion of the tasks that have been completed to date including a description of each of the issues to be addressed and the resolution of these concerns. Next, there is a presentation of the characteristics and demographics of the clients in the sample. Finally, a discussion of future work to be completed is included.

TASKS COMPLETED TO DATE

Several steps have been completed to reach the goal of validating New Mexico's RNA instrument. A review of the literature to assist decision-making regarding methodology has been completed. The following section explains the decisions made regarding the sample and variables collected. In addition, a description of the data collection process is included.

Sampling concerns

Closed versus open files; starting point and length of follow up

A fundamental issue that had to be addressed was deciding what constitutes a suitable sample.

The initial choice we faced was between sampling from active files or closed files. Our literature review indicated that closed files are most appropriate for a validation check. Although this may seem obvious, closed cases are chosen in order to guarantee that ample time has passed to measure recidivism.

A second concern that we had to resolve was when the follow up period should begin. There is some disagreement in the literature regarding at what point follow up should start. Bradshaw (1987) indicates that the follow up period should begin at the time of release from probation or parole. Conversely, Wagner (1991) defines follow up beginning from the offender's admission into probation or parole. Wagner's position is that an offender is at risk of failure from the moment he or she begins their probation or parole term, not from the time they finish their term. We defined a follow up period as beginning from the time a person exits probation or parole in order to measure long term outcomes.

This leads to the next concern: How long should the follow up period last? There seems to be a consensus that two years is minimum length for an adequate follow up period. If the follow up period is less than that, there is a greater likelihood of false positives. That is, that it is less likely that we will see recidivism among probation/parole clients because the follow up is too short to accurately measure recidivism. A follow up period of more than two years means that there is increased accuracy in measuring recidivism. As quoted in Bradshaw, Monahan (1978) suggests using multiple time periods for follow up. We primarily sampled from probation and parole cases that closed between January 1991 and June 1996. Thus, follow up periods range from two years to seven years. The January 1991 date was selected because that date corresponds to the earliest automated risk needs assessments. The end date was selected to ensure a minimum follow up period of two years. The analytical methods will account for the varying lengths of the follow up periods.

Sample size; sample representativeness

The natural successor to the question of what to sample is how many to sample. Sample size is an important issue, largely due to its impact on statistical inference and estimation. While we would like to be able to detect any actual effects of predictor variables in an analysis, it is quite possible to have a sample size too small to accomplish that goal. True effects can go undetected in small samples. A sample must be large enough to minimize this problem. The converse is true as well. That is, if a sample size is very large then nearly any effect-even those that are not

'real' effects-will be detected as statistically significance.

We handled the issue of sample size impact on detecting significant effects in two ways. The first way is to employ standard rules of thumb that generally provide a reasonable sample size to detect real effects without being so large as to lead to the detection of meaningless results. The most frequently utilized such rule is the ratio of the cases to the number of predictor variables in a statistical model. A lower limit for this ratio, suggested by Stevens (1995), is 10 cases per variable. Sometimes the suggestion is to use 30 cases per variable since the 'law of large numbers' generally takes effect in single variable samples of this size. We decided to use a ratio of 20 cases per variable, and expecting to explore statistical models with as many as 25 predictors in them, we arrived at a desired sample size of 500.

The second method of determining a sample size is by conducting a power analysis (Cohen, 1988). A power analysis determines how large a sample is needed to detect effects of a predetermined size, at a predetermined level of statistical significance, with a probability of 0.80 (or some other probability). As an example, we might deem it of interest to be able to detect any correlations between variables that was 0.30 or greater; the power analysis would inform us of the minimum sample size to achieve this outcome. A similar procedure can be followed for linear regression analysis. The power analysis was conducted, and the minimum sample size to detect statistical significance with a single predictor variable was 340. To deal with multiple predictor variables, we multiplied this sample size by the factor 1.2 to arrive at a final sample size of 408 files.

Both the rule of thumb and the power analysis are appropriate for statistical models when the dependent variable is normally distributed, as in the bell curve. Since some of the statistical models we expect to use are for categorical variables, the sample size was increased by 50% to account for the difference between normal variables and categorical ones. Assuming that we would inevitably have cases with missing data, we increased the sample size another 10% for this problem. Finally, we doubled the sample size to allow us to use half of the cases for evaluating the current RNA instrument and developing estimates, and the other half of the cases as a "hold-out sample" for validation of the estimates.

Following this procedure for the "rule of thumb" approach led to a projected sample size of 1650 ($=500*1.5*1.1*2$). The power analysis approach gave a sample size of 1346 ($=408*1.5*1.1*2$).

We settled on a final size of 2000 cases which exceeds the minimum projected sizes, and yet is not too large to detect false, or meaningless, effects. This also allows us to include more predictor variables in the model(s), if needed.

In order to accurately reflect the distribution of probation/parole clients throughout New Mexico, we counted the number of people in each region from a list of all probation and parole clients provided by the Probation and Parole Division. Then, we determined the number of files we would need from each region. Table 1 shows the breakdown of files needed from each of the four regions. The greatest proportion of cases is taken from Region II, which encompasses all of Bernalillo County and has the greatest number of clients.

Table 1. Files needed from each region

REGION	N	%
I	510	25.4
II	642	32.0
III	354	17.6
IV	502	25.0
TOTAL	2008	100

Variables

The selection of appropriate variables for the validation is essential. There are two types of variables that need to be selected: predictor and criterion variables. As the name implies, predictor variables are those that predict or explain an outcome. Criterion variables measure outcomes. Outcome measures include an individual's risk of failure on probation or parole and risk of future criminal behavior.

In addition to using risk and needs scores to predict a given outcome, other variables need to be included in order to determine whether they are better suited for prediction. The literature indicates that some variables consistently predict risk. These include personal characteristics such as age, gender, education, and whether a weapon was used during the current offense

(Bradshaw, 1987; Wright et al., 1984). Some variables have been found to have an inconsistent relationship with outcome measures. For example, Wright et al. (1984) found that address changes, prior convictions, prior probation periods, prior revocations, drug and alcohol use, and percent of time employed did not predict success among probationers. However, others report that these variables do predict outcomes (see Bradshaw, 1987). The additional predictor variables that we chose are those that 1) have been found to have a relationship with outcome measures; and 2) those we were able to collect because the information was consistently included in the files. Thus, some variables that may predict outcome, such as gang affiliation, were not collected because the information was not consistently recorded. We included the following variables as potential predictors in addition to the variables collected on the RNA: marital status, living arrangements, age of first offense, whether a weapon was used during the current offense, whether the client has spent any time in jail (excluding the current offense) since age fourteen, and amount of jail time served for current offense.

There are many types of criterion variables that can be chosen to assess outcomes. These can be separated into two types of outcomes: intermediate and long term. Intermediate outcomes include those that measure how a client performs while on probation or parole. Technical violations and how long a person was on probation or parole before technical violations occur are one way of measuring intermediate outcomes. Termination status, including whether the client absconded or was revoked, indicates performance on probation/parole. Additionally, whether the client tests positive for illicit substances, and how frequently, indicates a third type of intermediate outcome. Perhaps more important in regard to both overall success and risk to the community are the long term outcomes. These include whether the person was subsequently arrested and/or convicted and the length of time before the person was subsequently arrested/convicted. Both types of outcomes are important to include. It is possible that the RNA instrument currently in use will validly predict one type of outcome, but not another. Thus, the RNA may accurately predict how well a client behaves on probation/parole but does not predict recidivism. We have collected the intermediate outcomes from the case files. The long term outcomes, however, must be provided to us by Probation and Parole since we are not privy to the client's subsequent criminal behavior.

Data

Data was obtained from two sources. First, an automated database consisting of the risk and

needs scores from the initial assessment and final reassessment was provided by the Probation and Parole Division. Second, CARA staff collected hard copy data from closed probation and parole files.

Selection of files

Boxes of archived case files were obtained from Probation and Parole. Once the boxes were brought to the ISR, we chose which files to code based on the following criterion. First, files of clients who terminated probation or parole between January 1991 and June 1996 were targeted. Second, the RNA had to be either in the automated database already or the hard copy form was in the file and could be entered into the database. If the initial and final RNA forms were available, then we checked to make sure that none of the information was missing from the forms, including a check for out of range values. Third, as mentioned previously, we attempted to obtain a representative proportion of cases from each region. Once each of these items were checked for, a list of files to be coded was given to CARA research assistants.

Although we had established protocol for selecting files, there was some deviation from the protocol. For example, although the earliest end date was supposed to be January 1991, there were five that were prior to this date. Additionally, there were 47 clients that were released from probation/parole subsequent to June 1996. Despite the protocol violation, these were included because they provided greater representation from certain regions. However, most files (97.6%) were between our pre-specified dates.

Collection of data; accuracy of data collection

CARA staff coded information including demographics, intermediate outcomes, and prior and current offense information from the sampled Probation and Parole files. There were multiple potential sources of information in the files from which to gather the information needed. These included the following:

- Presentence Report (PSR). This was helpful for gathering information related to the client's prior record and demographics information.
- Order of Probation/Parole Certificate. This was used to gather information regarding the start date of probation or parole and terms of probation or parole including special conditions.
- Judgement and Sentence (J&S). This form contains information on the current charges the person was convicted of and the sentence given including any restitution ordered.

- Chronological Contact Sheets. These hand-written, sometimes typed, notes are sources of information for verifying the number of urinalysis test ordered and the number the client tested positive for, technical violations, initial living arrangements, and last date of probation or parole.
- NCIC Form, ACOPS and FBI Rap Sheet. These forms were used to collect information regarding the client's criminal history.
- Tox Screen reports. These forms provided information on the number of urinalyses tested, number positive, and type of substance for which a client tested positive.
- Report of Violation. This form contains information on probation or parole violations. Various court orders are included in the file that indicate the action taken as a result of the violation.
- Restitution Payments and receipts. These indicate the amount of restitution paid and still owed.

Additionally, the RNA instrument was used to verify information regarding start and end dates as well as income. Although income information is automated on the RNA form, the staff decided to collect the information from the hard copy records to ensure accuracy because the income from the automated database ranged from \$0 to \$7,530,000 per month.

We employed two methods of checking data collection accuracy. First, regular audits of coding were performed. Second, two inter-rater reliability checks were completed. The inter-rater reliability check consisted of choosing three types of files, one considered easy to code, one of some difficulty and one that was very difficult. Each of the research assistants were asked to code each of the files. The completed code books were examined by senior staff members for accuracy and consistency among the research assistants. Any problems noted were immediately addressed and rectified. Every effort was made to assure data accuracy and completeness. However, we were occasionally limited by the amount and quality of information available. These difficulties varied both across and within districts.

Some of the more frequent problems encountered noted by the research assistants are as follows. First, many files were not well organized despite a policy regarding the organization of files. This made it more difficult to find information. Second, many of the files did not contain all of the documentation. For example, many files were missing criminal histories, Judgement and Sentence forms, hard copy RNA instruments, and the test results from urine screens. Conversely, many of the same documents were duplicated two or three times throughout the

files. The forms most often duplicated were Orders of Probation, J&S, PSRs, and Violation Reports. Fourth, it was noted that the restitution records were often inaccurate. Fifth, not all probation or parole violations were noted either in the case notes or with a violation report. It was discovered that there had been a violation with references such as “previous inability to conform to the rules of probation.” Additionally, despite case notes that state that there was some violation or repeated violations, a violation report was not present. Sixth, it was found that when a client transferred supervision to another agency, the forms from the original file often did not appear in the new file. Each of these problems make it more difficult to find information, lead to missing data, and/or increase the likelihood of data collection error. Research assistants did note that certain regions and officers did an exceptional job at organizing files.

Data entry

There were 2201 cases collected and entered into a Microsoft Access database. Each of the cases were proofed for data entry errors. Once data entry and proofing was complete, this database was converted into an SPSS database. Cases that were collected by the ISR staff were matched by social security number and start date with those in the automated database provided by Probation and Parole. Sixty-five cases were dropped due to missing risk/needs assessment scores, because the dates between the automated database and the hard copy data did not correspond, or due to other missing information. This left 2136 cases in the resulting database.

DESCRIPTIVE INFORMATION

Characteristics of Probation/Parolees

In this section, we present descriptive data from the sample of closed files collected and coded by CARA staff. This descriptive data from the sample is supplemented, where possible, by summary descriptive information from the automated RNA database provided by the Probation and Parole Division (referred to as the population). This allows for a comparison between the sample files and the population of probation/parolees. Secondly, additional summary data coded directly from the sample of closed files is presented. This information primarily pertains to conduct while on probation or parole, and can be used to gain general insight into the clients the system is serving.

The characteristics of probation/parole clients is presented in Table 2. The first characteristic of

probationers/parolees presented is gender. Eighty-four-point-four percent (84.4%) of the 2136 usable cases are males. The proportion of males in the population is 84.0%, very close to the same level. The differences are not statistically significant, a finding that is somewhat surprising given the large number of cases (a total of 34,644) in the automated RNA database.

In terms of race/ethnicity, the breakdown in the sample and population are quite similar. Hispanics are slightly overrepresented in the sample (54.5%), compared to the population (50.8%). Whites are slightly underrepresented in the sample. These differences are statistically significant, but the overall profiles of race/ethnicity between the sample and population are quite similar and not meaningfully significant. Recall from the discussion of sample size above, if the sample size gets big enough, any differences will be statistically significant, even those differences which are not very meaningful.

The distribution of age at intake is quite similar for the sample and the population, as is the mean age at intake (31.08 for the sample, 31.28 for the population). Neither differences are statistically significant. The largest portion of probation/parolees are 24 years old or younger (31.1% in the sample; 28.2% in the population).

Education in the sample and the population are strikingly similar. No statistically significant difference between the sample and the population can be detected. Marital status cannot be compared between the sample and the population as marital status is not included in the automated RNA database. Among the sample, most clients are not married; only 21.5% are currently married, and almost half (47.2%) have never been married.

Employment is distributed roughly the same in the sample and the population, however there are statistically significant differences. There are fewer full-time employed clients (35.5%) in the sample than in the population (39.0%), and more unemployed but looking clients in the sample.

Examining the households of the sample, we find that the vast majority of clients in the sample (71.3%) are living with a spouse or partner, or with parents. The distribution in the population is unavailable as this information is not part of the automated RNA database. Note, also, that the percentages of the sample add up to more than 100% as a client could be living with people that come from more than one of the categories. Client household comparisons for number of dependents is quite similar between the sample and population. Nearly half the clients in the

sample (45.4%) and the population (48.2%) have no dependents. There is no statistically significant difference between the sample and population when it comes to number of dependents.

The vast majority of the sample and population are on probation (82.9% of the sample, 81.2% of the population). Slightly more of the clients in the population are on parole (17.5%) compared to those in the sample (15.5%). However, this difference is not statistically significant.

Sample and population differences in level of supervision at intake are not statistically significant. Nearly half the clients are placed in maximum supervision (48.5% for the sample, 49.6% for the population). However, the level of supervision at discharge does show some differences. More of the sample clients are at minimum security, and slightly fewer are at maximum. Initial Risk Assessment Scores are equivalently distributed in both the sample and the population. Initial Needs assessment is likewise similarly distributed in both the sample and the population.

Finally, termination type is statistically different between the sample and the population. The difference is essentially attributable to cases that terminate by being transferred out of state. Only 0.6% of the sample have been transferred, but 5.4% of the population have. Other than this large difference, the proportions in the sample roughly match the proportion in the population. Half of all clients (51.5% in the sample; 48.8% in the population) have successful terminations (either satisfactory or early termination). Only a very small fraction (0.5% for the sample; 0.4% for the population) terminate as absconders. About one-fifth of all clients (21.8% of the sample; 22.7% of the population) terminate in revocation.

Table 2.

Characteristics of Evaluation/Validation Sample and Total Probation and Parole Population

Probationer/Parolee Characteristic	Sample Cases	% Sample	% Total Population
Gender			
Male	1811	84.8%	84.0%
Female	325	15.2%	16.0%
Total	2136	100.0%	100.0%
Race/Ethnicity			
White	658	30.8%	33.6%

Probationer/Parolee Characteristic	Sample Cases	% Sample	% Total Population
Black	124	5.8%	6.9%
Hispanic	1164	54.5%	50.8%
Native American	167	7.8%	7.6%
Other	23	1.1%	1.1%
Total	2136	100.0%	100.0%
Age at Intake:			
Up to 24	662	31.1%	28.2%
25 to 29	413	19.4%	20.7%
30 to 34	388	18.2%	19.5%
35 to 39	295	13.8%	14.3%
40 and above	372	17.5%	17.3%
Total	2130	100.0%	100.0%
Mean Age (standard deviation)		31.08 (9.88)	31.28 (9.54)
Education:			
0 to 8 years	202	9.5%	9.5%
9 to 12 years (or GED)	1535	71.9%	70.7%
13 years or more	399	18.7%	19.8%
Total	2136	100.0%	100.0%
Marital Status:			
Single, never married	980	47.2%	N/A
Married	447	21.5%	
Divorced	419	20.2%	
Separated	98	4.7%	
Widowed	35	1.7%	
Common law marriage	96	4.6%	
Total	2075	100.00%	
Employment:			
Full Time	759	35.5%	39.0%
Part Time	188	8.8%	8.3%

Probationer/Parolee Characteristic	Sample Cases	% Sample	% Total Population
Unemployed (looking)	724	33.9%	30.6%
Unemployed (not looking)	286	13.4%	12.7%
Seasonal	22	1.0%	1.1%
Student	50	2.3%	2.7%
Disabled	107	5.0%	5.5%
Total	2136	100.0%	100.0%
Living Arrangements at Intake:			
Spouse/Partner	803	41.2%	N/A
Parents	587	30.1%	
Children	424	21.8%	
Other Family	241	12.4%	
Self Only	218	11.2%	
Friends	93	4.8%	
Other	74	3.8%	
Total	1949 clients	125.3%	
Number of Dependents:			
None	969	45.4%	48.2%
One	430	20.1%	18.4%
Two	307	14.4%	14.8%
Three	234	11.0%	10.2%
Four	127	6.0%	5.2%
Five	36	1.7%	2.0%
Six or more	31	1.5%	1.2%
Total	2134	100.0%	100.0%
Mean		1.24	1.17
(standard deviation)		(1.53)	(1.47)
Age at 1st Offense:			
Mean	1906	22.23	N/A
(standard deviation)		(8.98)	

Probationer/Parolee Characteristic	Sample Cases	% Sample	% Total Population
Program type			
Probation	1769	82.9%	81.2%
Parole	331	15.5%	17.5%
Reintegration	34	1.6%	1.4%
Total	2134	100.0%	100.0%

Probationer/Parolee Characteristic	Sample Cases	% Sample	% Total Population
Level of Supervision (at intake):			
Minimum	140	6.6%	6.6%
Medium	835	39.1%	37.0%
Maximum	1035	48.5%	49.6%
Intensive	126	5.9%	6.8%
Total	2136	100.0%	100.0%
Level of Supervision (at discharge):			
Minimum	598	28.2%	25.94%
Medium	784	36.9%	35.24%
Maximum	634	29.9%	31.73%
Intensive	106	5.0%	7.10%
Total	2122	100.0%	100.00%
Initial Risk Assessment Score:			
Low	517	24.2%	25.7%
Medium	522	24.4%	23.1%
High	1096	51.3%	51.2%
Total	2135	100.0%	100.0%
Initial Needs Assessment Score:			
Low	282	13.2%	15.7%
Medium	1109	51.9%	52.3%
High	745	34.9%	32.0%
Total	2136	100.0%	100.0%
Final Risk Assessment Score:			
Low	1019	47.9%	50.8%
Medium	443	20.8%	19.4%
High	666	31.3%	29.8%
Total	2128	100.0%	100.0%

Probationer/Parolee Characteristic	Sample Cases	% Sample	% Total Population
Final Needs Assessment Score:			
Low	813	38.4%	42.3%
Medium	697	32.9%	29.9%
High	608	28.7%	27.8%
Total	2118	100.0%	100.0%
Termination Type:			
Satisfactory	941	44.1%	40.8%
Unsatisfactory	519	24.3%	21.8%
Revoked	464	21.8%	22.7%
Early	158	7.4%	8.0%
Transfer out of state	13	0.6%	5.4%
Death	27	1.3%	1.0%
Absconder	11	0.5%	0.4%
Total	2133	100.0%	100.0%

The next table, Table 3, provides descriptive information on the sample of 2136 cases only. This is data that was coded from the archived files provided by the Probation and Parole Division. Given the strong similarity between the sample and the population presented in Table 2, we can presume that the distribution of characteristics in the Table 3 for the sample would be quite similar to the population, though we cannot test this hypothesis with an actual comparison.

The current offense of the sample of probation/parolees is the initial characteristic displayed. Note that probation/parolees can be admitted with multiple charges, meaning that the percentages will add up to greater than one-hundred percent. The most common charges are drug offenses (22.6%), DUI (13.4%), and burglary/breaking-and-entering (13.2%). Among violent offenses, murder and rape are poorly represented perhaps reflecting that persons committing those crimes tend to serve longer prison sentences and are less likely to fall into the sample of probation/parolees. A substantial percentage (17.5%) of offenses in the sample are categorized as "all other offenses." These include charges such as conspiracy, various tampering charges (i.e., tampering with evidence and tampering with a motor vehicle), trespassing, habitual

offender, and contributing to the delinquency of a minor.

A vast majority of the probation/parolees were incarcerated for the current offense (79.2%). The average jail time is 7.89 months. Quite likely, a large portion of the jail time was time prior to conviction and sentencing in light of the fact that most of the population is on probation. Only slightly over half (55.2%) the sample had been incarcerated since age fourteen.

Turning to characteristics of the sample that pertain to performance while on probation or parole, we first see that a large portion of the probation/parolees committed technical violations (61.3%). Of those that committed technical violations, they averaged 4.23 violations per client. Less than half the sample (43.3%) was required to undergo any urinalysis. Most clients (65.8%) who had one or more urinalyses completed tested positive at least once. The remaining 34.2% never tested positive. Of the 901 clients that did have to undergo urinalysis, there were a total of 3787 performed (an average of 4.20 per client). Of the total number of urinalyses completed, well over half (60.7%) were clean. Only 38.7% of the tests turned up some substance. The most common such substances were cannabis (51.6%) and cocaine (28.1%). Note that the percentages of substances adds to greater than 100% since a single urinalysis could turn up multiple substances.

Table 3. Conduct of Probationer and Parolees in Evaluation/Validation Sample

	Sample Cases	% Sample
Current Offense		
Drug Offenses	484	22.6%
Driving under the influence	287	13.4%
Burglary-Breaking and Entering	282	13.2%
Aggravated Assault	192	9.0%
Larceny-Theft	185	8.7%
Traffic Offenses	123	5.8%
Stolen Property (buying, receiving, possessing)	120	5.6%
Offenses against a Police Officer	116	5.4%
Fraud	103	4.8%
Other Assault	89	4.2%

	Sample Cases	% Sample
Forgery and Counterfeiting	88	4.1%
Motor Vehicle Theft	72	3.4%
Robbery	65	3.0%
Vandalism	56	2.6%
Probation/Parole Violation	55	2.6%
Offenses against Family and Children	55	2.6%
Weapons (buying, possessing, etc.)	51	2.4%
Sex Offenses (not including rape and prostitution)	46	2.2%
Embezzlement	44	2.1%
Murder	40	1.9%
Rape	36	1.7%
False Imprisonment	26	1.2%
Failure to Appear/Bench Warrant	25	1.2%
Disorderly Conduct	21	1.0%
Liquor Law Violations	16	0.7%
Arson	10	0.5%
Drunkenness	1	0.0%
All Other Offenses	373	17.5%
Total	2136	100.0%
Incarcerated for Current Offense:		
Yes (Mean Jail Time of 7.89 months)	1589	79.2%
No	418	20.8%
Total	2007	100.0%
Incarcerated Since Age 14 (excluding current offense):		
Yes	1009	55.2%
No	820	44.8%
Total	1829	100.0%

	Sample Cases	% Sample
Restitution:		
Ordered to Pay Restitution	603	29.1%
Restitution Waived	(4)	(0.7%)
Paid all of Restitution	(245)	(40.6%)
Paid Part of Restitution	(212)	(35.2%)
Did Not Pay Any Restitution	(107)	(17.7%)
Missing Information on Payment	(35)	(5.8%)
Not Ordered to Pay Restitution	1466	70.9%
Total	2069	100.0%
Technical Violations:		
Yes (Mean of 4.23 technical violations)	1307	61.3%
No	825	38.7%
Total	2132	100.0%
Urinalysis Taken:		
Yes, at least one	901	43.3%
None	1181	56.7%
Total	2082	100.0%
Urinalyses Results		
At Least One Positive result	587	65.8%
Never positive	305	34.2%
Total	892	100.0%
Urinalysis Results of Total Number Taken		
Total Number Clean	2300	60.7%
Total Number Dirty	1466	38.7%
Missing Information	21	0.6%
Total	3787	100.0%

	Sample Cases	% Sample
Substances for which Clients Tested Positive		
Cannabis	756	51.6%
Cocaine	412	28.1%
Opiates	248	16.9%
Amphetamines	215	14.7%
Alcohol	151	10.3%
Benzodiazepines	57	3.9%
Methadone	18	1.2%
Barbiturates	8	0.5%
Phencyclidine	6	0.4%
Proxyphene	4	0.3%
Total	1466	100.0%

Summary of comparison of sample to population

The overall comparison of clients in the sample with clients from the population is quite positive. Most of the comparisons indicate general similarity between the sample and the population. While several comparisons resulted in statistically different distributions, the detection of such small differences is really simply due to the large population sample.

Summary of characteristics of sample

Most of the clients are on probation rather than parole or reintegration. The level of supervision at intake for both clients in the sample and population is higher than at discharge. Likewise, a greater percentage of clients have lower risk and needs scores over time. Offenses related to drugs (distribution/possession) or alcohol (DUI) are among the most prevalent in the sample. Non-violent offenses outnumber violent offenses. Most clients in the sample spent some time incarcerated due to the current offense. A slight majority had been incarcerated prior to this offense. Most probationers and parolees did violate the terms of their probation/parole at least once. However, most of the clients were discharged from probation or parole rather than revoked or absconded.

FUTURE TASKS

There are several goals we wish to accomplish during the next funding cycle. First, the validation check needs to be completed. Second, based on the results of the validation check, any revisions deemed necessary will be recommended. Third, a revised instrument, if necessary, will be created. Fourth, implementation of the revised instrument should occur. Fifth, an inter-rater reliability check among the probation and parole officers using the revised instrument is necessary. A validation check on the revised instrument should follow. However, this cannot happen until the revised instrument has been in use long enough to get a large enough sample of closed cases in which the new instrument was administered. Thus, it is not feasible for this last goal to occur during the next funding cycle (July, 1998 to June, 1999). None of these tasks can be completed until we receive the subsequent arrest and conviction data on the sample of clients that has been drawn. Thus far, it has been decided that for clients who have a felony conviction, we will be able to get subsequent arrest/conviction information from the FBI. However, for those who are misdemeanants, and therefore do not have an FBI number, the source of their subsequent arrests and convictions is not yet clear. Unfortunately, a large proportion of the sample (21%) are misdemeanants. Without the long term outcomes for this portion of the sample, the RNA instrument cannot be validated.

Validation

Data entry of subsequent criminal behavior

Once we have the subsequent arrest and conviction information, we can proceed with the validation check. We will first enter subsequent arrests and convictions directly from the hard copy information that we receive into a database.¹ This is likely to be a time-consuming process. In order to facilitate the process, we will have a database that contains the client's name and social security number as identifiers and the beginning date of supervision. This way, the data entry person can simply look up a client and enter their subsequent offenses. Assuming that it would take 15 minutes per case, with three staff working 15 hours per week, it will take approximately 12 weeks to complete this task.

¹Subsequent arrests and convictions are defined as any arrests and/or convictions that occur after the client began probation or parole.

Data analysis

Once the entry of subsequent offenses is completed, we will then analyze the data. The analysis of the probation/parole sample of closed files will proceed with several stages: evaluation of current instrument, revision of current instrument, and validating revisions.

Evaluating the current instrument. The first stage is to evaluate the current RNA form. This is accomplished by determining how well the current scoring system discriminates between success and failure of a probationer/parolee. Success and failure will be ascertained with several different outcome measures. Conduct while on probation or parole is one measure of success or failure. The number of technical violations and the speed at which a client commits the initial technical violation, as well as the tendency for a probationer/parolee to have negative urinalysis tests are indicative of progress or the lack of progress. Additionally, the type of termination (successful, absconder, etc.) from probation/parole is an indicator of success. Finally, the timing and frequency of subsequent arrests is the strongest indicator of success or failure. One of the data analysis techniques that will be used for this step is correlation coefficients to establish whether there is a linear relationship between the scores and the outcomes (for example, are the scores increasing with increasing subsequent arrests?). Second, for binary (any subsequent arrests or not) and categorical outcomes, we will use contingency tables.

Revising the current instrument. This evaluation of the current RNA scoring system will suggest avenues for improving the RNA form and scoring system. Thus, the second task is to develop a revised RNA form that better discriminates successful probationers/parolees from the unsuccessful. All the outcomes discussed above will come into play, though the most important outcome is subsequent criminal behavior. Various statistical techniques will be considered to determine the variables that should be included in the revised RNA form, as well as the weighting scheme for scoring that form. These techniques include Poisson regression for count outcomes (such as number of subsequent arrests), logistic regression for binary outcomes (whether there are any subsequent arrests), multinomial regression for categorical outcomes (such as termination status), and survival models for outcomes involving the timing of recidivism (length of time before the first arrest). Several specifications of each model will be explored to determine the predictors that result in optimal discrimination of success and failure. Half of the 2136 cases will be used for RNA evaluation.

Final validation. After this stage, the new RNA and scoring system will be created, and then

Implementation of revised instrument and inter-rater reliability check

Once the instrument is revised and approved by all involved, the instrument will be implemented. This means that all those who will administer the new instrument must be trained. We will assist in this training in conjunction with Department of Corrections staff. Additionally, a check of inter-rater reliability among those filling out the forms will be completed. The details regarding the logistics of completing phase, including how long it is expected to take, will have to be determined in the future.

Once the data regarding recidivism is received, it will take approximately seven and one-half months to validate, write a report (with recommendations) and revise the instrument. The remainder of the funding period can be used to implement a revised instrument. However, again, this is all dependent on the receipt of subsequent offenses for our sample.

Bibliography

Bradshaw, Richard Alfred. Multivariate Actuarial Prediction of Felonious Recidivism of Male Parolees: Development and Cross-Validation of a Series of Risk Assessment Models Using Stepwise Logistic Regression. University Microfilms International: Ann Arbor. 1987.

Clear, Todd. "Statistical Prediction in Corrections". *Research in Corrections*, 1:1, March, 1988.

Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences* (revised edition). Academic Press: New York, 1977.

Gendreau, Paul. American Probation and Parole Association. Targeting an Appropriate ISP Population. Chapter II-2, Risk/Need Assessment.

Institute for Social Research, University of New Mexico. "Risk and Needs Assessment in Sentencing Decisions: A Summary Report". Working Paper No. 5. March, 1996.

Stevens, Joseph. Class lecture on multivariate techniques. Education Foundations 603: Multivariate Analysis, University of New Mexico. Fall, 1995.

Wright, Kevin N., Todd R. Clear and Paul Dickson. "Universal Applicability of Probation Risk-Assessment Instruments". *Criminology*, 22:1, February, 1984.