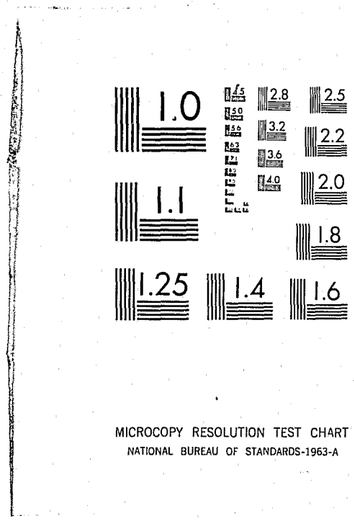


MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

National Institute of Justice
United States Department of Justice
Washington, D. C. 20531

DATE FILMED

1-26-82

MF-1

FINAL REPORT (# AP-0)

MICHIGAN RISK PREDICTION:

A REPLICATION STUDY

Terrence H. Murphy
Michigan Dept. of Corrections
Program Bureau
Lansing, Michigan

This study was prepared under Grant Number AP-0 from the National Institute of Corrections, Bureau of Prisons, U.S. Department of Justice. Points of view or opinions stated in this document are those of the author and do not necessarily represent the official position or policies of the U. S. Department of Justice.

The original risk study, "Predicting the Relative Recidivism Risk of Parolees," was prepared under Grant Number 76-ED-05-0017 from the Law Enforcement Assistance Administration, U. S. Department of Justice.

THE PROBLEM AND RELATED ISSUES

Statement of the Problem

The problem of trying to determine who will commit violent and nonviolent crime while on parole has been a major focus of the Michigan Department of Corrections. Building upon the findings of previous parole prediction studies, a major study was undertaken to isolate characteristics which identify subgroups within the parolee population which have either a higher or lower than average probability of committing serious crime on parole.

The results of the original risk study found characteristics identifying different subgroups which have probabilities of committing new violent crime on parole that ranged from a high of 40% to a low of 2% (contrasted to the base rate for the entire group of 10.5%). The probabilities of committing new nonviolent crime on parole ranged from a high of 37.9% to a low of 15.1% (contrasted to a base rate of 25%).

The purpose of this study is to see if the findings of the initial risk study can be replicated by conducting the same analysis on another random sample of parolees.

Review of Original Risk Study

The population examined in the original study consisted of all male inmates paroled between January, 1971 and December, 1971 in the State of Michigan. The population excluded females because their number was so small (3% of the total) that multivariate analysis would not be productive.

A randomly selected 50% sample of 1971 parolees (N = 2200) was chosen. This sample size seemed the best compromise given the data collection restric-

tions, time constraints, desire for representativeness and frequency requirements for certain variables (e.g., homicide is a relatively infrequent crime and a large sample is necessary to obtain sufficient numbers of homicides to permit analyses).

Data were collected on 360 variables per individual. The independent variables were categorized into three different time frames: Time 1, any variable or measure which could generally be related to the inmate prior to his time of incarceration for the instant offense; Time 2, any variable which would be appropriately measured during the inmate's incarceration for the instant offense; Time 3, any variable or measure which generally related to the inmate's parole.[1] The dependent variable, recidivism, measured the actual behavior of the crime rather than administrative-legal decisions affecting a person's parole completion. A five-point recidivism scale was devised for use in the study. The scale categories were: no illegal activity, technical violation, misdemeanor, nonviolent felony and violent felony. Only the most serious applicable category was coded in each case.

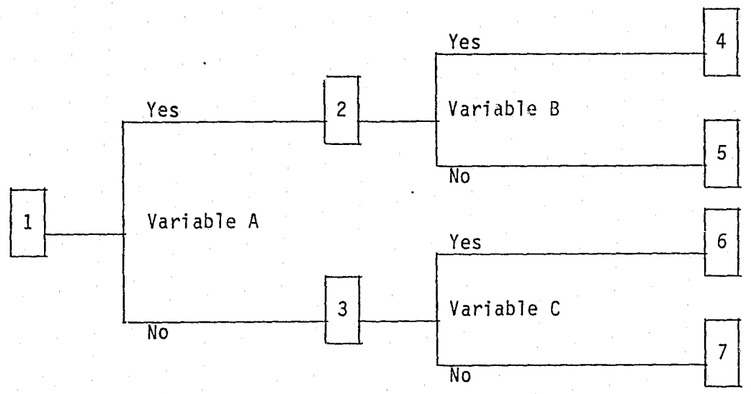
The final risk groups were developed through the use of AID (Automatic Interaction Detection).[2] The AID program, developed at the University of Michigan, is a technique designed to detect significant interactions among a large number of independent variables (predictors) in relation to a single dependent or criterion variable. The AID analysis was used to systematically search and select from many possible combinations of predictors those which produced the highest and lowest rate of violent crime.

---

[1] A more detailed discussion of the original study and its application may be found in the Michigan Department of Corrections publication "Information on Michigan Department of Corrections' Risk Screening," August 15, 1978.

[2] For more specific details on AID, the reader is referred to John Sonquist, et al. Searching for Structure, University of Michigan, Ann Arbor, 1973.

The following diagram outlines the procedure illustratively:



In the original study the analysis produced the following groups with their respective rate of new violent felony on parole:

RISK GROUP	RECIDIVISM RATE*	PERCENT OF SAMPLE
Very High Risk. Instant offense of rape, robbery, or homicide and serious misconduct or security segregation and first arrest before 15th birthday.	40.0%	4.7%
High Risk. Instant offense of rape, robbery, or homicide and serious misconduct and age of first arrest was 15 or over.	20.7%	6.6%
Middle Risk. Instant offense either rape, robbery, or homicide and no serious misconduct; or instant offense not rape, robbery or homicide and reported felony while juvenile.	11.83%	45.5%
Low Risk. Instant offense not rape, robbery, or homicide (may be other assaultive crime) and no reported felony while juvenile and never been married at time of instant offense.	6.3%	23.5%
Very Low Risk. Instant offense not rape, robbery, or homicide and no reported felony while juvenile and not serving on other assaultive crime and has been married.	2%	19.7%
*Base Rate	10.5%	

This same search and select procedure was used to identify different configurations which differentiated the rates of nonviolent felony on parole. The dependent variable was dichotomized to nonviolent felony, "yes" or "no." Those who had committed a violent felony on parole were excluded from this analysis because they would have automatically been put in the "no nonviolent felony" group in the dichotomization. This would have suppressed, to some extent, the ability to differentiate the rates of nonviolent felony. The configural analysis produced the following groups and their respective rates of new nonviolent crime on parole:

RISK GROUP	RECIDIVISM RATE*	PERCENT OF SAMPLE
High Risk. Reported felony while juvenile and major misconduct; or reported felony while juvenile and no major misconduct and age of first arrest before 15th birthday.	37.9%	28.2%
Middle Risk. Reported felony while juvenile and no major misconduct and first arrest after 15th birthday; or no reported felony while juvenile and drug problem at the time of instant offense.	27%	28.4%
Low Risk. No reported felony while juvenile and no drug problem at time of instant offense.	15.1%	43.4%

*Base Rate 25%. This rate was calculated by first excluding all those who had committed a violent felony on parole, thus reducing the sample size to N = 1820. The base rate of the nonviolent felony for the entire sample was 22.5%.

These configurations represent subgroups of the paroling population. For any particular prisoner to be considered a member of any one of these subgroups, he must have all of the characteristics which define that group since the respective subgroups are defined by the combined interactions of the variable characteristics. Recidivism rates for the subgroups refer to the mean (average) rate of violent or nonviolent felony on parole for the

respective subgroups. These results do not mean that it is possible to specify recidivism for a particular individual but rather indicate that he is a member of a particular subgroup which had a rate of recidivism that, in the case of the very high risk group for example, was nearly four times as high as the base rate.

Methodological Notes on Both Studies

Several aspects of both the original and replication studies deserve individual attention. The purpose of this section is to briefly examine those aspects of the research process that may be considered relevant to both studies and/or improvements over previous attempts at risk prediction.

1. Sample Size - Both studies utilize large samples. The traditional reliance on smaller samples is questionable when a population is to be split into successively smaller groups through the analysis of many variables with a variety of "cutting scores." Much past research has suffered from samples too small to permit subgroup analysis. Glaser (1964) recognized small sample size as a major limitation of prediction tables in general. Also, Sonquist, et. al., has specifically warned users of AID that:

> data sets with a thousand cases or more are necessary; otherwise the power of the search processes must be restricted drastically or those processes will carry one into a never-never land of idiosyncratic results. (1973:3)

Finally, the need for large sample sizes is often influenced by the occurrence of the phenomenon (dependent variable) under study. In this case, the base rate[3] of violent crime is so low (10.5%) that a small sample would restrict predictability.

---

[3] "The 'base rate' refers to the proportion of individuals in some population who fall into a category which is to be predicted." (Gottfredson, 1970:752). For example, 10.5% of the 1971 sample committed a violent felony (category to be predicted).

2. Random Sampling Procedures - Few prediction studies have used random samples from the population of interest.

> Frequently, a prediction method is devised on the basis of study of a sample containing equal (or about equal) numbers of delinquents and non-delinquents. Then it is applied to the general population where the proportion who actually become delinquent is considerably lower. This procedure can be expected to result in a serious overestimation of the practical effectiveness of the prediction device.
> (Gottfredson, 1970:753)

In this case, both studies are based on large random samples of parolees. As stated earlier, the sampling procedure reflects a concern over time constraints on the one hand and a desire for representativeness and frequency requirements on the other.

3. Follow-up Period - Many studies of parole prediction and recidivism have been criticized because of their short follow-up period and/or because of their exclusive reliance on crime measures defined in the context of a parole period.

In both the original and replication studies, the period of follow-up consisted of the time a resident was on parole during the three-year period (36 months) immediately subsequent to his parole for the instant offense. The mean length of parole during that period was approximately 14 months.

As in other studies, the measure of recidivism (defined in the section below) was defined in the context of the parole period. The rationale for this is closely related to the desire to have the recidivism measure reflect actual overt behavior to the greatest extent possible. This required the utilization of detailed descriptions of criminal behavior

from both parole agent investigations and police reports. Such detail is not often found in other measures of recidivism nor available when an individual is no longer under parole supervision.

4. Dependent Variable; Recidivism - One of the many ways in which these studies differed from most previous research in risk prediction is that they sought to measure actual criminal behavior rather than judicial or administrative decisions (which may or may not follow from that behavior) affecting a person's parole completion.

Traditionally, analyses of recidivism have used a dichotomous dependent measure -- success/failure on parole. The definition of success/failure often obscures the actual behavior of interest. Such an either-or categorization fails to differentiate degrees of behavior on parole. For example, it may be inappropriate to equate return to prison because of a new felony such as homicide or kidnapping with return to prison due to a technical violation. These measures may fail to account for felonies receiving probation; technical violations of a criminal nature v. technical violations of parole conditions only; and the length of time necessary for criminal dispositions. Consequently, a dependent variable expressed in terms of success/failure on parole frequently makes the dependent variable a function of administrative and policy conditions affecting parole continuance rather than actual human behavior. The following recidivism scale, used in both studies, was designed to overcome these limitations:

1. No Illegal Activity
2. Technical Violation
3. Misdemeanor
4. Nonviolent Felony
5. Violent Felony

This scale is a behavioral index of the inmate's most serious behavior while on parole. For example, if a parolee had only committed a minor technical violation on parole and nothing else, he would have been coded a two on this scale. If he had committed both a misdemeanor (three) and a violent felony (five) while on parole, he would have received a "five." The coding criteria for this scale was based upon written descriptions from police and agent records of the behavior involved in the criminal activity. The criteria did not rest upon arraignments nor convictions but tried to reflect as closely as possible the actual reported behavioral description of the man's activity. This resulted from the department's major focus on preventing felonious behavior on parole with particular attention directed towards violence.

5. Data Collection - The primary source of data for both studies was inmate files which included Central Office files, institutional files, and parole files for each resident. Data collection and coding was conducted by college graduates who were hired and trained specifically for these projects. Separate office space was rented for both studies to minimize distractions.

   The reliability of the coding was established in the following manner: Approximately 20 inmate files were randomly selected from the samples and the variables were coded by the coders hired for the project. A reliability coefficient was calculated for each of the variables as well as for each of the coders. If the reliability coefficient was less than .90 the variable was either dropped from the variable list or revised so that it could be coded with a .90 reliability. This means that if any of the given variables were not coded the same way by the coders, it

indicated that either the variable was ambiguous or that there were problems with missing data. In either case, the variable was usually dropped from further consideration. Similarly, a reliability coefficient of .90 was established among all of the coders. People not coding information reliably were replaced.

6. Violence Prediction - The Michigan Department of Corrections had greatest concern for felony behavior on parole. Although tables projecting both nonviolent and violent felony risk are used for classification and release decisions, violence receives the primary emphasis. Few studies exist in this specific area and the successful identification of any subgroups in this regard deserves careful attention. Although the statistical findings will show the extent of false positives and false negatives, the acceptable level of these errors for practical purposes remains an administrative decision. The focus on felony behavior provides a more specific criteria for such decisions in terms of public protection and administrative benefits.

DESIGN OF THE REPLICATION STUDY

Population and Sample

The population considered in this replication study consisted of all inmates in Michigan paroled from January 1, 1974 through December 31, 1974. The sample represents 30% of that population or 1,200 subjects.

A computer printout of all persons paroled in 1974 was obtained from the Bureau of Management Services of the Michigan Department of Corrections. All persons on the printout were sequentially numbered after which a random numbers table was used to draw the sample.

Data Collection

Data were collected on 90 variables per individual. The magnitude of the task necessitated strict supervision of coding, clear specification of variables, and adequate coder training. As described earlier, coder and variable reliability was established at the 90% level. The entire coding process lasted seven months.

Variables

The variables were categorized into the following groups:

1. Descriptive Variables: These included basic demographic and background characteristics for between sample comparisons and within sample descriptions (e.g., race, community size).

2. Exploratory Variables: Certain variables were added because of problems encountered in the original study. For example, during the original

study it was suggested that some parolees raised outstate may have incomplete criminal histories. We wanted to examine the possible influence of this artifact.

3. Independent/Dependent Variables: The criterion variable (recidivism) and the independent variables which defined the risk groups in the 1971 sample.

4. Refinement Variables: These items were included for further analyses. Some had shown significant correlations in the original study but were not as powerful as those used in the risk prediction tables. Also included within this category were a number of parole variables for evaluating parole performance and stability.

Analysis

The analysis was conducted in three phases. The first phase involved the creation of a computer program to duplicate the subgroups of the original study. The SPSS: Statistical Package for the Social Sciences (Nye, et al, 1975) was used for group formations and a variety of statistical comparisons. The violent and nonviolent felony rates were then determined for each of the respective subgroups in the 1974 sample.

The second phase addressed the replication issue. The replication findings were examined from the following perspectives.[4]

---

[4] Validation and/or replication is seldom discussed in terms of the conceptual issues involved. Efficiency, effectiveness, stability, and comparability (between samples) are a few of the different perspectives taken.

For an example of contrasting views, the reader is referred to A.J. Kahn's criticism in "The Case of the Premature Claims," Crime and Delinquency, 11, 217-228, 1965 and P.G. Ward's "Validating Prediction Scales" Brit. J. of Criminology, 1967, 7:136:44. Because of the numerous and diverse approaches, this project focused on those issues most consistent with the department's goals and concerns.

1. Do the subgroups in the replication study result in the predicted order of very low to very high risk? A few authors have allowed for the precise order of the groups to change between samples yet still be considered valid (Simon, 1971). For practical use in classification, the replication study would have failed if high risk cases have a lower probability of committing a felony than low risk cases.

2. The groups were examined internally using only the replication data to see if the expected order holds up and the subgroups other than middle risk are statistically distinct from the base rate. Given that most actuarial, AID and clustering programs focus on identifying subgroups of homogeneous subjects that are substantially different from the rest of the sample, the base rate of the criterion variable is an important element. For instance, the published discussion concerning the original violent risk groups had said that "the results of this risk study identified characteristics of different subgroups which have probabilities of committing new violent crime on parole that ranged from a high of 40% to a low of 2% contrasted to the base rate of 10.5%." The department was interested in the extent to which this statement remained true.

3. A popular approach to evaluting validation or replication results is comparing the Mean Cost Ratings (MCR) for each sample (Simon; 1971, Glaser, 1964). The MCR is one of a series of measures developed by Duncan, et al (1951) for use in making selection decisions. The MCR varies between 0 (no differentiation) to 1. An advantage of the MCR over other measures is that it is sensitive to the order of the risk groups. As Simon points out, the MCR "is influenced not only by the degree of separation achieved between the risk groups, but also by the

extent to which they are in strict order of failure rates." (1971:21) Consequently, the measure is consistent with the department's concern in (1) above. It is important to recognize that due to regression and differences between samples, some shrinkage in group probabilities is to be expected. The primary concern is evaluating the amount of shrinkage. Changes in group characteristics over time may account for some of the observed differences.

4. Related to the above perspective, the department was interested in the differences between the 1971 sample and the 1974 sample in terms of: a) differences in failure rates between subgroups and, b) differences in the distribution of cases among the subgroups, since the proportion of cases within each subgroup has practical implications for classification purposes. Simon has noted that "shrinkage may be caused by a change in class failure rates, or a change in the distribution of cases through the various classes, or both of these things." (1971:28).

The final phase of the analysis was concerned with exploring reverse validation procedures with AID as well as examining the impact of selected control variables on parole outcome. A series of variables were substituted in the replication analysis to determine if stronger discriminators existed that continued to discriminate when applied to the original 1971 sample. In addition, exploratory analysis focused on controlling for the effects of certain variables in an attempt to specify the relationship between risk level and felony behavior.

FINDINGS AND DISCUSSION

Descriptive Comparisons Between Samples

The samples were compared on basic demographic and institutional characteristics. The results are presented in table 1. Although the comparisons show some differences, the samples are quite similar on most of the variables considered.

Of particular interest are those variables that are (or have been) risk predictors. The juvenile variables (age at first arrest, commission of an adult felony while a juvenile, juvenile commitments) show no significant differences between 1971 and 1974. These variables have traditionally been related to future criminal activity. The samples are also similar on the variables 'marital status' and 'raised by both parents.' 'Age at start of prison' for the instant offense and the proportion of first offenders has also remained consistent as has the critical institutional behavior variable of segregation.

While the samples are similar in terms of the proportion of first offenders, the 1974 sample has slightly more subjects with prior prison commitments. The same contrast occurs with the length of time incarcerated. Although the average age at the start of their prison term remained constant, the length of stay increased for the replication subjects. Other contrasts between samples include racial composition (the 1971 sample had fewer minorities), and alcohol and drug problems.

These comparions, albeit brief, show the samples to be similar with respect to the majority of variables examined. More importantly, the samples show

Table 1

COMPARISON OF SAMPLE CHARACTERISTICS (1971 v. 1974)

Variable	Original	Replication	Result	Probability
Age at first arrest				
$\bar{X}$	17.44	17.49	t = .257	.80>p>.70
s.d.	5.4	5.0		
Adult felony while juvenile				
yes	870(43%)	471(39.9%)	X^2=3.78	.10>p>.05
no	1134	710		
Juvenile commitment				
yes	396(19.6%)	254(21.4%)	X^2=1.42	.30>p>.20
no	1620	933		
Race				
white	974(47.9%)	513(42.5%)	X^2=7.63	.01>p>.001
non-white	1059	683		
Raised by both parents				
yes	1211(60%)	702(60.1%)	X^2= .004	.95>p>.90
no	808	466(39.9%)		
Marital status				
single	1117(60%)	679(57.5%)	X^2=1.91	.20>p>.10
non-single	915	502		
Alcohol problem				
yes	635(31%)	306(26.4%)	X^2=8.92	.01>p>.001
no	1384	852		
Drug problem				
yes	479(24%)	375(32.5%)	X^2=28.62	p<.001
no	1539	779		
Prior prison commitment				
yes	676(33.2%)	445(36.8%)	X^2=4.15	.05>p>.01
no	1357	765		
Prior felonies				
yes	1299(64%)	755(62%)	X^2=0.76	.50>p>.30
no	733	455		
Age at start of prison				
$\bar{X}$	25.4	25.6	t = 0.1	p=.90
s.d.	8.1	8.2		

Table 1 (con't.)
Comparison of Sample Characteristics (1971 v. 1974)

Variable	Original	Replication	Result	Probability *
Number of months incarcerated				
$\bar{X}$	26.8	30.4	t = 4.12	p<.001
s.d.	21.2	29.2		
Involuntary segregation				
$\bar{X}$	12.13	11.96	t = .095	p>.90
s.d.	50.54	48.18		

*The probabilities were evaluated against a .05 level of significance.

no differences on those variables found to be predictive of future criminal activity in the original study (with the exception of drug problem).[5]

Base Rate Contrasts

The major purpose of the original risk study was to identify subgroups within the parolee population which have either a higher or lower than average probability of committing violent or nonviolent crime on parole.

Violent felony on parole. Table 2 classifies the replication sample into the subgroups defined in the original study. The probabilities of committing a violent felony on parole as found in the replication study are shown for each group. The original order (rank) was found to be retained. The probability of that happening if in fact no relationship existed with the dependent variable is 1/120 (assuming distinct outcomes). The table also compares each subgroup with the base rate. It was expected that VLR and LR would be significantly lower; middle risk would show no significant differences; and HR and VHR would be significantly higher than the base rate.

All of these expectations were confirmed. Table 2 shows that the study replicated the exact order of the subgroups with significantly different probabilities of committing a new violent felony on parole when contrasted with the base rate. The two extreme groups were regressed toward the mean, however, and the "very low risk" group in particular changed from a 2% to a 9% failure rate. Note also that the sample base rate is 16%, compared with 10.5% in the original study. These issues will be discussed further in the section on between group comparisons.

[5] It should be noted that these comparisons are provided for descriptive purposes only. Because the AID results are based on interactions of variables, one should not infer a relationship between a specific variable and a particular subgroup.

Table 2

COMPARISON OF VIOLENT RISK GROUPS WITH BASE RATE

RISK GROUP	SUBGROUP RECIDIVISM RATE	REPLICATION BASE RATE	Z*	PROBABILITY
Very High Risk. Instant offense of rape, robbery, or homicide and serious misconduct or security segregation and first arrest before 15th birthday.	.320	.16	3.09	(p = .001)
High Risk. Instant offense of rape, robbery, or homicide and serious misconduct and age of first arrest was 15 or over.	.279	.16	3.01	(p = .001)
Middle Risk. Instant offense either rape, robbery, or homicide and no serious misconduct; or instant offense not rape, robbery, or homicide and reported felony while juvenile.	.174	.16	.99	(N.S.)**
Low Risk. Instant offense not rape, robbery, or homicide (may be other assaultive crime) and no reported felony while juvenile and never been married at time of instant offense.	.111	.16	2.09	(p = .018)
Very Low Risk. Instant offense not rape, robbery, or homicide and no reported felony while juvenile and not serving on other assaultive crime and had been married.	.089	.16	2.76	(p = .003)

*Difference of proportions, one-tail test at .05.
**Two-tailed test.

Property Felony on Parole. The same information regarding property felonies is presented in Table 3. It was expected that LR would be significantly lower; middle risk would show no significant differences; and HR would be significantly higher than the base rate. The hypotheses are supported. In addition to replicating the exact order, low risk and high risk groups commit significantly different probabilities of property felonies when contrasted with the base rate.

Mean Cost Rating (M.C.R.)

The M.C.R. is one of several approaches to evaluating the shrinkage between samples in addition to providing an estimate of a classification device's power.

Violent Felony Behavior. The results for both samples with respect to violent crime on parole are presented in Table 4. An examination of the 1971 sample shows the overal difference between failure rates to be significant (as measured by X^2). When calculated for the original sample, M.C.R. = .39.

Similar results are found in the 1974 replication sample. The predicted order of the risk groups in terms of failure rates is retained under both conditions. The overall difference in failure rates continues to be significant (as measured by X^2). However, the replication sample, with an M.C.R. of .23, indicates that a shrinkage in power does occur over this time period. As one would expect, the shrinkage shows the tendency of outlying risk groups to gravitate toward the mean. In addition, more cases are classified as middle risk. Both of these factors contribute to the reduction in power observed between samples.

Table 3

COMPARISON OF PROPERTY RISK GROUPS WITH BASE RATE

RISK GROUP	SUBGROUP RECIDIVISM RATE	REPLICATION BASE RATE	Z*	PROBABILITY
High Risk. Reported felony while juvenile and major misconduct; or reported felony while juvenile and no major misconduct and age of first arrest before 15th birthday.	.340	.258	3.15	(p = .0008)
Middle Risk. Reported felony while juvenile and no major misconduct and age of first arrest 15 or over; or no reported felony while juvenile and drug problem at the time of instant offense.	.298	.258	1.54	(N.S.)**
Low Risk. No reported felony while juvenile and no drug problem at time of instant offense.	.174	.258	4.0	(p .0001)

*Difference of proportions; one-tail test at .05.
**Two-tailed test.

Table 4

VIOLENT RISK PREDICTION TABLE:
Original and Replication

1971 Sample

Risk Group	Success	Failure	TOTAL	Failure Rate (%)	Percent of Total in Class
Very Low	397	8	405	2.0	19.9
Low	405	29	434	6.7	21.3
Middle	843	107	950	11.3	46.8
High	116	29	145	20.0	7.1
Very High	59	40	99	40.4	4.9
TOTAL	1,820	213	2,033	10.5	100.0

X^2 = 146.6 df. = 4 Sign. = .0000
M.C.R. = .39

1974 Sample

Risk Group	Success	Failure	TOTAL	Failure Rate (%)	Percent of Total in Class
Very Low	185	18	203	8.9	17.2
Low	217	27	244	11.1	20.6
Middle	495	104	599	17.4	50.7
High	62	24	86	27.9	7.3
Very High	34	16	50	32.0	4.2
TOTAL	993	189	1,182	16.0	100.0

X^2 = 31.6 df. = 4 Sign. = .0000
M.C.R. = .23

It has already been emphasized that shrinkage is to be expected. It may be worth noting that traditional recidivism devices based on the broader criterion of success/failure (without regard to the specific behavior or violence) usually attain an M.C.R. of .12 to .31 upon validation with a substantial amount of shrinkage (Simon, 1972). Given the additional consideration of different time periods, the replication sample compares favorably with these efforts. However, the results do indicate that the replication sample (1974) has considerably less power than the original sample (1971).

Property Felony Behavior. The chi squares and Mean Cost Ratings for property recidivism are presented in Table 5. The results are similar to the findings on violence in that the property risk groups continue to be significant upon replication (as measured by X^2). The original risk results have an M.C.R. of .27 as compared to the replication's M.C.R. of .21. The degree of shrinkage experienced with the violence tables does not occur in the case of property comparisons. The tables indicate a minimal amount of shrinkage (.06). The similarity between years is, in part, due to the smaller amount of discrimination in the original table.

Between Group Contrasts

The final issue concerning replication was the identification of any differences between samples with respect to failure rates and the distribution of cases across subgroups.

Distribution of Sample Cases. The subgroup comparisons were conducted through a series of chi-square and difference of proportions tests. Each risk level (i.e., very low, low) in the 1971 sample was compared with it's respective level in the 1974 sample.

TABLE 5

PROPERTY RISK PREDICTION TABLE:
Original and Replication

1971 Sample

Risk Group	Success	Failure	TOTAL	Failure Rate (%)	Percent of Total in Class
Low	670	119	789	15.1	43.4
Middle	378	139	517	26.9	28.4
High	319	195	514	37.9	28.2
TOTAL	1,367	453	1,820	24.9	100.0

Base rate excluding violent = 24.9
X^2 = 88.5 2 df. Sign. = .0000
M.C.R. = .27

1974 Sample

Risk Group	Success	Failure	TOTAL	Failure Rate (%)	Percent of Total in Class
Low	347	73	420	17.4	42.3
Middle	200	85	285	29.8	28.7
High	190	98	288	34.0	29.0
TOTAL	737	256	993	25.8	100.0

Base rate excluding violent = 25.9
X^2 = 28.16 Sign. = .0000
M.C.R. = .21

Table 6 addresses whether the proportion of cases within subgroups changed between samples. The results indicate that the distribution of subjects among violent risk levels remained stable. There are no significant differences at any level between years. The largest observed difference (though not statistically significant) occurs in the middle risk group for violent offenders (4%). The property felony contrasts also show a remarkable consistency in the distribution of cases across subgroups.

Comparison of Failure Rates. Whether the samples differ in failure rates is examined in Table 7. The failure rates for low, high and very high violent risk groups are not statistically different. A comparison of property felony rates also shows no significant differences between samples on any level of risk. The only significant differences between years occur in the very low and middle risk groups for violence.

The observation that the very low and middle violent risk groups differ in failure rates between years suggests that these samples are not from the same population. Because the distribution of cases among subgroups remained relatively consistent over the years, the source of these differences is difficult to specify. Among the possible explanations for the differences are:

a) Marital status or its combination with other factors does not discriminate as strongly as it did in the past. Given the rapid changes in the social significance of marital status, the differences may not be surprising. The effect of excluding marital status as a criterion for subgroup membership is illustrated in diagram 1. The procedure combines single and non-single offenders

Table 6
COMPARISON OF RISK GROUP MEMBERSHIP BETWEEN SAMPLES
Violence

RISK LEVEL		SAMPLE 1971	SAMPLE 1974	Z*	PROBABILITY
Very Low	N	398	203		
	Prop.	.196	.172	1.69	p = .091
Low	N	441	244		
	Prop.	.217	.206	1.0	p = .317
Middle	N	950	599		
	Prop.	.468	.507	2.14	p = .032
High	N	145	86		
	Prop.	.071	.073	2.13	p = .834
Very High	N	99	50		
	Prop.	.049	.042	.921	p = .358

Property

RISK LEVEL		SAMPLE 1971	SAMPLE 1974	Z*	PROBABILITY
Low	N	789	420		
	Prop.	.434	.423	.564	p = .5754
Middle	N	517	285		
	Prop.	.284	.287	.169	p = .865
High	N	514	288		
	Prop.	.282	.29	.452	p = .6528

*The criterion for rejection in tables 6 and 7 is based on a series of two-tailed tests at the .01 level of significance. It was noted that the middle risk group would have been significant at the .05 level. As stated earlier in this report, scant information is available concerning replication efforts in prediction. A series of meetings were conducted prior to the analysis stage and the issue of replication focused on a series of two sample tests between the original data and replication data using the .01 level of significance. However, because the middle risk group is closely related to the base rate, a few comments regarding the differences in base rates and middle risk groups are found later in this report.

Table 7

COMPARISON OF FAILURE RATES
BETWEEN ORIGINAL AND REPLICATION SUBGROUPS

Violence

Subgroup Comparison		Difference of Proportion	Significance*
Very Low	v. Very Low	4.12	p .0001
Low	v. Low	2.02	p = .044
Middle	v. Middle	3.40	p .001
High	v. High	1.38	p = .168
Very High	v. Very High	1.0	p = .317

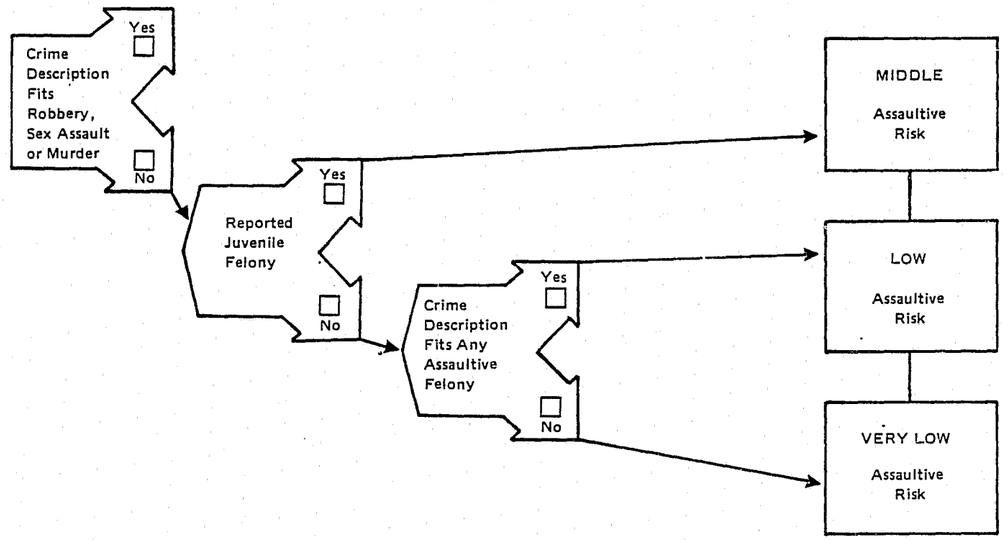
Property

Subgroup Comparison		Difference of Proportion	Significance*
Low	v. Low	1.04	p = .298
Middle	v. Middle	.879	p = .378
High	v. High	1.09	p = .275

* Because of the number of two sample tests involved, the .01 level of significance was used. It was noted that the low risk group would have been significant at the .05. While the following discussion focuses on those differences found to be significant, the low risk group receives considerable implication because of its adjacent position to very low risk.

DIAGRAM 1

THE PROCESS OF EXCLUDING MARITAL STATUS
AS A CRITERION OF SUBGROUP MEMBERSHIP



into the very low risk group. The data (see Table 8) suggests that marital status does not discriminate between groups. A comparison of Table 4 (Pg. 20) with Table 8 indicates that marital status tends to suppress the violence rate for those classified as low risk in the 1974 sample. The impact on the proportion of cases lower than the base rate is minimal (37% vs. 33%).

b) Circumstances may have changed that possibly affect the reliability of certain variables. Because of the growing concern over the use and confidentiality of juvenile histories, it is possible that an increased number of cases are missing this data (falsely shown as having no such history). For this reason, the replication study added a new variable, "raised outstate." It was suggested that those offenders raised in other states, particularly in rural areas, may have past histories unavailable to Michigan investigators. If that is the case, excluding that group of offenders should improve the discrimination among groups. When this is done as in table 9, we find support for this hypothesis. The impact of excluding "raised outstate" on the proportional distribution of cases is not very pronounced for any risk group. The largest observed difference occurs in the very low risk group (2%). The exclusion of these cases also refines the failure rates for both of the extreme groups (very low, very high). Both of these groups require juvenile data. Table 9 suggests that with improved juvenile data especially for those raised outstate, it may be possible to discriminate among all five groups.[6]

[6] The basis for examining outstate cases was the possible impact of missing data on the results. For purposes of analysis, multiple controls in conjunction with 'outstate' did not appear meaningful at this point.

Table 8

VIOLENT RISK GROUPS WITH MARITAL STATUS EXCLUDED

Risk Group	Success	Failure	TOTAL	Failure Rate (%)	Percent of Total in Class
Very Low	358	34	392	8.7	33.2
Low	44	11	55	20.0	4.7
Middle	495	104	599	17.4	50.7
High	62	24	86	27.9	7.3
Very High	34	16	50	32.0	4.2
TOTAL	993	189	1,182	16.0	100.0

Table 9

VIOLENT RISK LEVELS EXCLUDING OUTSTATE CASES

(1974)

Risk Group	Success	Failure	TOTAL	Failure Rate (%) Outstate Excluded	Failure Rate (%) (Entire Sample)	Percent of Total in Class Outstate Excluded	Percent of Total in Class (Entire Sample)
Very Low	136	10	146	6.8	(8.9)	15.2	(17.2)
Low	183	23	206	11.2	(11.1)	21.5	(20.6)
Middle	411	89	500	17.8	(17.4)	52.1	(50.7)
High	47	17	64	26.6	(27.9)	6.7	(7.3)
Very High	28	15	43	34.9	(32.0)	4.5	(4.2)
TOTAL	805	154	959	16.1		100.0	

c) It is also possible that the increase in violence for the total sample may, for reasons unknown, have been particularly aggravated among those cases identified as very low risk. 1974 was marked by a recession (high unemployment) and nationwide increases in the violent crime rate (U.C.R., 1975). It should be noted that the sample years and respective follow-up periods (1971-1976) have in general shown an increase in assaultive prison commitments. The question which is raised is whether the impact of unstable social conditions may be strongest upon those people not likely otherwise to become involved in violence.

The increase in the violence base rate is directly related to, and may entirely explain, the differences observed between middle risk groups. As stated earlier, the middle risk group was not expected to differ from the sample base rate. Part of the increases in the base rate and middle risk group may also be related to the increase in the number of persons classified as middle risk. As noted earlier, the largest observed difference (though not statistically significant) occurs in the middle risk group for violent offenders. The next largest difference occurs in the very low risk group where 2.4% fewer people are classified as such. While these differences are not statistically significant, it is felt that they should be noted. Generally, the fewer persons classified as middle will increase the power of the table.

Each of the above are possible explanations for the differences observed between samples. At this point they remain speculative.

Discussion

To a large degree, the present study replicates the original risk results. The findings, particularly the sections on base rate contrasts and mean cost ratings support this conclusion. At the same time, the between group contrasts and the shrinkage in mean cost ratings for the violence table show the violatile nature of any prediction device that attempts to differentiate groups with extreme probabilities. If viewed as an artifact of statistical regression, the more extreme a subgroup rate differs from the mean rate, the more likely that rate will move toward the mean rate when measured again. Whether the amount of shrinkage for the very low risk group is indicative of the unusual social conditions of 1974, the characteristics that define this group or a product of data-availability, requires additional research if a definitive answer is to be obtained. The property tables, on the other hand, show no significant differences between years with regard to failure rates or case distribution.

Some exploratory analysis has focused on substituting selected variables for marital status to increase the discrimination between low and very low risk groups.[7] Whether the results are artifacts of the 1974 sample can only be addressed in future comparisons. Another area of interest is the relationship between recidivism and parole related variables.[8] Preliminary analysis of employment variables indicates that the 1971 sample had more subjects employed full-time than the 1974 sample at the time of their parole

[7] The selection process controlled for the effects of additional variables on subgroup probabilities with marital status excluded.

[8] As stated earlier, one phase of the analysis was concerned with exploring reverse validation procedures in an attempt to identify stronger discriminators in the 1974 sample that continued to discriminate on the 1971 sample. At this point, efforts using AID and other approaches have not produced results that are better than those produced by the original study. These efforts have been directed towards entering a series of new variables and the replacement of 'key split' variables.

termination (45% vs. 31% respectively). In both studies, the variable final job status had significant zero-order correlations with recidivism (1971 = -.33; 1974 = -.29). An examination of employment patterns on parole outcome controlling for risk level may provide additional information on the role of employment on parole adjustment.

A final conclusion for the replication results is the need for continued validation in later years. The benefits of further validation exercises are twofold. First, the need for revision of risk parameters and/or characteristics is well documented. Second, a later sample would provide insight into the impact of unstable social conditions upon risk groups. It is quite possible that a later sample may show probabilities different from those attained in the original and replication samples.

REFERENCES

Babst, Dean V., Don M. Gottfredson and Kelley B. Ballard, Jr.

1968 "Comparison of multiple regression and cluster analysis techniques for developing base expectancy tables." Journal of Research in Crime and Delinquency 5(1):72-80.

Blalock, Hurbert M.

1972 Social Statistics, 2nd Ed. New York: McGraw-Hill

Duncan, O.D., L. E. Ohlin, A. J. Reiss, Jr., and H. R. Stanton

1953 "Formal devices for making selection decisions." American Journal of Sociology 58:573-584.

Glaser, Daniel

1955 "The efficiency of alternative approaches to parole prediction." American Sociological Review 20:283-287.

1964 The Effectiveness of a Prison and Parole System. New York:Bobbs-Merrill.

Gottfredson, Don M.

1970 "Assessment of prediction methods." Pp. 745-771 in Norman Johnston, Leonard Savitz, Marvin E. Wolfgang (eds.), The Sociology of Punishment and Corrections. New York:Wiley.

Henkel, Ramon E.

1976 Tests of Significance. Sage University Paper on Quatitative Applications, 07-044. Beverly Hills:Sage.

Kahn, Alfred J.

1965 "The case of the premature claims." Crime and Delinquency 11:217-228.

Nie, Norman H., C. Hadlai Hyll, Jean C. Jenkins, Karin Steinbrenner and Dale H. Bent

1975 SPSS: Statistical Package for the Social Sciences: 2nd Ed. New York:McGraw Hill.

Simon, Frances, H.

1971 Prediction Methods in Criminology. London: Her Majesty's Stationary Office.

1972 "Statistical Methods of Making Prediction Instruments." Journal of Research in Crime and Deliquency 9:46-53.

Sonquist, John A., Elizabeth L. Baker, James N. Morgon.

1973 Searching for Structure. (Rev. Ed.) Ann Arbor: Institute for Social Research, University of Michigan.

Ward, P.G.

1970 "Validating Prediction Scales." Pp. 801-806 in Norman Johnston, et al (eds.), The Sociology of Punishment and Corrections. New York:Wiley.

END