Washington State Institute for Public Policy

110 Fifth Avenue SE, Suite 214 • PO Box 40999 • Olympia, WA 98504-0999 • (360) 586-2677 • FAX (360) 586-2793 • www.wsipp.wa.gov

January 2003

WASHINGTON'S OFFENDER ACCOUNTABILITY ACT:

UPDATE AND PROGRESS REPORT ON THE ACT'S EVALUATION

In 1999, the Washington Legislature passed, and Governor Locke signed into law, the Offender Accountability Act (OAA). The Act primarily affects how the state provides community supervision to adults convicted of felony crimes.

In terms of broad policy direction, the OAA amended state law by adding a seventh goal to Washington's sentencing policy. The six previously established goals pertain to ensuring that sentences are proportionate with the severity of an offender's past and current crimes; that sentences are commensurate with those imposed on others convicted of similar offenses; and that the sentencing system uses taxpayer resources frugally.

The new seventh goal establishes that reducing the "risk of reoffending by offenders in the community" is also a purpose of Washington's sentencing policy.[1] By adding this new policy directive, the Legislature indicated that an explicit purpose of Washington's laws is not only to sentence people according to the amount of harm they have done in the *past*, but also to reduce the *future* rate of re-offending once convicted offenders are back in the community.

To implement this policy, the OAA directs the Washington State Department of Corrections (DOC) to:

a) Classify felony offenders according to the risk they pose to re-offending in the future, and the amount of harm they have caused society in the past, and

---

[1] The current statute (RCW 9.94A.010) now reads that the seven purposes of Washington's sentencing laws are to:
(1) Ensure that the punishment for a criminal offense is proportionate to the seriousness of the offense and the offender's criminal history;
(2) Promote respect for the law by providing punishment which is just;
(3) Be commensurate with the punishment imposed on others committing similar offenses;
(4) Protect the public;
(5) Offer the offender an opportunity to improve him or herself;
(6) Make frugal use of the state's and local governments' resources; and
(7) Reduce the risk of reoffending by offenders in the community.

Summary

In 1999, the Washington Legislature passed the Offender Accountability Act (OAA), and full implementation of the Act began by 2001. The Act affects how the state provides community supervision to adults convicted of felony crimes. The OAA directs the Department of Corrections (DOC) to classify all felony offenders according to the risk they pose to re-offending in the future, and the amount of harm they have caused society in the past. The OAA then directs DOC to allocate more of its community-based resources to the higher-risk offenders. The primary goal is to reduce the subsequent criminal behavior of these offenders when they are back in the community.

The OAA directs the Washington State Institute for Public Policy to conduct an evaluation to determine if the Act achieves reduced re-offense rates and improvements in other outcomes. The Institute must report annually on the evaluation.

It is too early in the life of the OAA to determine if the Act has had an effect on recidivism rates. Because a sufficient follow-up period is needed to observe recidivism, our January 2005 report will offer the first opportunity to test whether the OAA reduces crime cost-effectively.

In this year's report, we present the first results on how well DOC's risk assessment tool—the Level of Service Inventory-Revised (LSI-R)—predicts actual recidivism. We examine 27,288 offenders assessed with the LSI-R from 1998 to October 2000 (before the OAA went into effect) and calculate how many were re-convicted for another crime after they were back in the community for 12 months. We find that the LSI-R is able to predict subsequent criminal behavior reasonably well. Those offenders with higher LSI-R scores have considerably higher 12-month recidivism rates than those with lower LSI-R scores.

In this report, we also describe some of the technical statistical steps we are taking to ensure that the OAA's outcomes can be reliably evaluated.

b) Deploy more resources to higher-risk offenders and, with a relatively fixed budget, spend correspondingly fewer dollars on lower-risk offenders.[2]

The legislature intended that the changes initiated with the OAA would reduce overall re-offense rates in Washington. To see if this happens, the OAA directs the Washington State Institute for Public Policy (Institute) to evaluate the Act's effect on re-offense rates (often referred to as "recidivism") and several other outcomes described in the legislation.[3] The Institute is directed to prepare annual reports on the evaluation, with a final report due January 1, 2010. This report is this year's installment.

The principal question for the evaluation of the OAA is whether the Act produces a positive *net* benefit for Washington. The word "net" is emphasized because the OAA involves a fundamental tradeoff. With relatively fixed funding levels, if more DOC resources are devoted to higher-risk offenders, then fewer resources will be available for lower-risk offenders. If supervision and treatment resources are efficacious then, under the OAA, recidivism rates can be expected to go down for higher-risk offenders and up for lower-risk offenders.[4] Therefore, the ultimate question for the evaluation of the OAA is whether the gains will outweigh the losses. That is, will Washington obtain a *net* benefit with the OAA?

It is too early in the life of the OAA to answer this question. The first substantial group of offenders classified by DOC pursuant to the OAA occurred during calendar year 2001. We will be able to track this group's recidivism rate only after they have been released from prison or jail. Some higher-risk offenders convicted of serious crimes receive prison terms that extend for many years. It is necessary to wait until these offenders are released to the community before recidivism rates can be observed. Additionally, once offenders are released from incarceration, a sufficient "at-risk" follow-up period must be allowed to test whether the OAA reduces recidivism.

For adult offenders, we have found that the minimum "at-risk period" needs to be at least 24 months, although preliminary results can be calculated with 12- and 18-month follow-up periods. We have also found that a 12-month "adjudication period" is required to assure that any re-convictions are determined and recorded by the court in Washington's court-based information system. This is why it can take three years from the time a program starts before the recidivism outcomes of a policy change, such as the OAA, can be reliably determined.[5] Thus, it is too early in the Institute's evaluation of the OAA to determine if it is reducing recidivism rates and if the Act's gains outweigh the losses. Because of these necessary follow-up and adjudication periods, our January 2005 report will offer the first opportunity to test preliminarily whether the OAA reduces crime cost-effectively.

This Year's Report. In this year's OAA report, we present information on two topics. First, we show the first statistical results on the accuracy of the "risk assessment" tool adopted by the DOC. As we discuss, the OAA requires DOC to use a formal research-based screening tool to assess which offenders are most likely to re-offend. We present new results on how well DOC's chosen instrument predicts actual recidivism rates for the first substantial group of offenders assessed by DOC.

The second topic covered is a progress report on how we are creating the "comparison group" that the Institute will ultimately use to evaluate whether the OAA lowers the overall recidivism rate and affects other outcomes. Any evaluation of outcomes is only meaningful if it can answer the question: Compared to what? The Legislature wanted to know whether the OAA produces better results when compared with the system in place prior to the OAA. Therefore, in our evaluation of the OAA, we will compare those offenders affected by DOC's implementation of the OAA, which began in 2001, with a similar group of offenders in the pre-OAA period. In this report, we describe the technical statistical procedures we are using to select the comparison group.

---

[2] The OAA also gave DOC new authority to hold timely hearings and to sanction offenders who violate conditions of community custody for offenders with crimes committed after July 1, 2000.

[3] RCW 72.09.610.

[4] If, however, community supervision and treatment resources do not have an effect on recidivism rates, then, of course, the OAA will not change the recidivism rates of either the higher-risk or the lower-risk offenders. In the evaluation of the OAA, the Institute will be able to test for this question.

[5] All Institute reports follow the statistical definition for recidivism that the 1997 Legislature directed the Institute to establish. See: R. Barnoski (1997) *Standards for Improving Research Effectiveness in Adult and Juvenile Justice*, Olympia: Washington State Institute for Public Policy.

Part One: DOC's Risk Management Identification System and Resource Allocation Under the OAA

As described, the OAA involves two broad steps: (a) DOC must classify all offenders using a research-based assessment tool, and (b) DOC must then supervise and treat offenders in accordance with the results of the classification.

The OAA instructs DOC to classify felony offenders according to the risk they pose to re-offending in the future, and the amount of harm they have caused society in the past. To give operational direction to this new policy, the OAA defines risk assessment with this language:

> *"Risk assessment" means the application of an objective instrument supported by research and adopted by the department for the purpose of assessing an offender's **risk of reoffense,** taking into consideration the **nature of the harm done** by the offender, place and circumstances of the offender related to risk, the offender's relationship to any victim, and any information provided to the department by victims.[6]*

With this language, the Legislature indicated that it wanted DOC to classify offenders by taking into account two broad concepts: the "risk of re-offense" and the "nature of the harm done." These two concepts do not necessarily address the same thing.

The "risk of re-offense" concept is forward looking. A classification system that measures the risk of re-offense is designed to predict whether an offender is likely to commit another crime in the future. The "harm-done" concept, on the other hand, is backward looking. A classification system that measures harm done, measures how much damage an offender has already caused victims and society, regardless of what he or she is likely to do in the future.

DOC designed its "Risk Management Identification" (RMI) system to include two sets of assessments and decision rules that, together, attempt to measure and balance both of these OAA concepts. First, DOC adopted a formal risk assessment tool to measure the likelihood of future re-offending. Second, DOC adopted

additional criteria to gauge how much harm the offender's prior criminal activity caused victims and society. Each of these two tools is summarized here.[7]

1. DOC's "Risk of Re-Offense" Assessment Tool. Prior to the OAA, DOC began using a formal risk assessment tool called the "Level of Service Inventory-Revised (LSI-R)." Canadian researchers developed this 54-question, copyrighted instrument in the 1980s. There is some previous research (not done in Washington) indicating that the LSI-R is a valid way to predict whether an offender is likely to re-offend.[8] DOC adopted the LSI-R as one of the key parts of its Risk Management Identification system.

The 54 questions on the LSI-R cover ten areas of an offender's life. These include: ten questions on prior criminal history; ten questions on an offender's education and employment; two questions on finances; four questions on an offender's family situation; three questions on an offender's living situation; two questions on leisure and recreation activities; five questions on peers; nine questions on alcohol and drug problems; five questions on emotional or personal problems; and four questions on an offender's attitude.

After DOC staff administers the LSI-R, an offender's combined LSI-R score is tabulated. An offender's LSI-R score can range from 1 to 54, where higher numbers indicate a higher probability of re-offending.

In Part Two of this report, we present an analysis that shows the degree to which the LSI-R predicts actual recidivism rates in Washington.

2. DOC's "Harm-Done" Criteria. The LSI-R was designed to predict whether an offender will commit another crime in the future. It was not, however, constructed to measure the level of prior harm caused by an offender—a key requirement

---

[6] RCW 9.94A.030 (35), emphasis added.

[7] In 2002, the Institute's annual OAA report evaluated how well DOC's adopted classification system identifies offenders based on the likelihood of re-offense, and the nature of the harm done by the OAA offenders. We found that DOC's system does a reasonable job classifying offenders pursuant to the policy directives of the OAA. See: S. Aos (2002) *Washington's Offender Accountability Act: An Evaluation of the Department of Corrections' Risk Management Identification System*, Olympia: Washington State Institute for Public Policy; available at: <www.wsipp.wa.gov/crime/pdf/OAA_Jan02.pdf>

[8] Prior research associated with the LSI-R is discussed in: D. A. Andrews and J. L. Bonta (1995) *The Level of Service Inventory-Revised, Manual*. North Tonawanda, New York: Multi-Health Systems, Inc.

in the OAA legislation. To implement this aspect of the OAA, DOC adopted an additional set of rules to gauge how much damage an offender has caused in his or her prior criminality. DOC developed these harm-done rules from recommendations by DOC staff, the Victims Council, and criteria established by the Washington Association of Sheriffs and Police Chiefs.[9] Examples of these rules include the following: Is the offender classified as a Level I, II, or III sex offender? Is the offender designated as a Dangerous Mentally Ill Offender? Did the offender commit a violent offense involving a stranger? If an offender scores a "yes" on these conditions, then—regardless of the offender's LSI-R score—the offender is regarded as needing higher levels of community custody.[10]

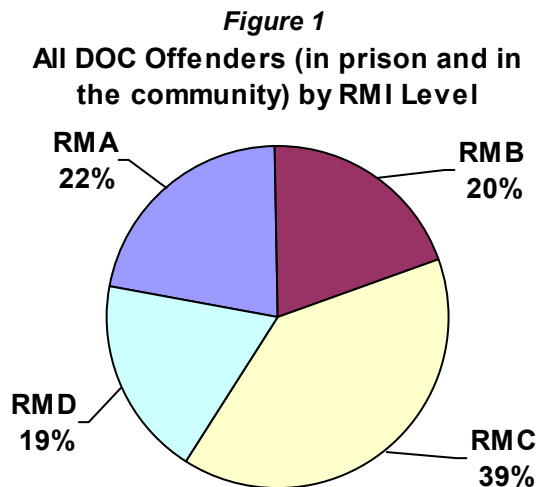**The Product of the RMI System: RMA, RMB, RMC, and RMD Offender Classifications.** Together, the LSI-R and the harm-done criteria make up DOC's RMI classification system, and DOC uses the system to classify each offender. There is also an "override" procedure in the RMI system where an offender can be moved among the RMI levels if a DOC officer deems it necessary and receives supervisory approval for the re-classification. With these scoring rules, each felon under DOC supervision is classified as one of the following: an RMA, RMB, RMC, or RMD offender. The RMA category is the highest-risk and harm-done classification, while the RMD category is the lowest-risk group. Figure 1 shows recent data on the distribution all RMI-classified DOC offenders (both those in prison and those in the community). Forty-two percent of offenders are classified as either RMA or RMB (the highest-risk and harm-done categories), with the majority (58 percent) of DOC offenders classified in the less risky RMC and RMD categories.

Why Is the RMI Designation Important? Answer: DOC Resource Allocation. These classifications are significant because the OAA directs DOC to deploy the bulk of its community-based resources to the higher-risk RMA and RMB offenders, with correspondingly fewer resources devoted to the relatively lower-risk RMC and RMD offenders. Thus, the RMI designation is central to the OAA and whether the Act is able to reduce overall recidivism rates.

---

[9] Washington State Department of Corrections, *Risk Assessment and the Offender Accountability Act, November 5, 2001,* presented to the House Criminal Justice and Corrections Committee, November 30, 2001.
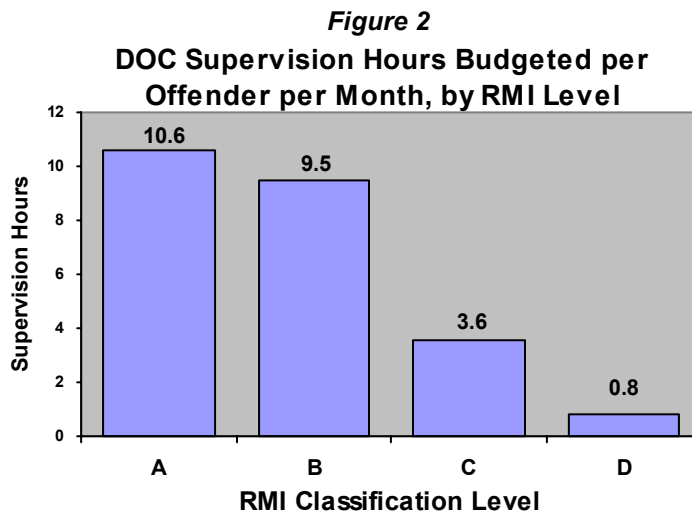[10] The Appendix to the Institute's 2002 OAA report contains a copy of DOC's RMI Worksheet. See: S. Aos (2002).

Figure 1
All DOC Offenders (in prison and in the community) by RMI Level

Source: Institute analysis of data in: DOC Briefing Document, House Criminal Justice and Corrections Committee, December 6, 2002.

Recent DOC data indicate that the high-risk (and high harm-done) RMA offenders receive an average 10.6 hours of community supervision per month by DOC staff. The slightly lower-risk RMB offenders receive a similar average level of supervision: 9.5 hours of staff attention per month. RMC and RMD offenders, on the other hand, receive considerably less supervision at 3.6 hours and .8 hours, respectively. These numbers are shown on Figure 2.

Thus, these DOC data provide a clear indication that more resources are being spent on the higher-risk offenders and correspondingly fewer resources are being spent on lower-risk offenders. Whether the OAA is able to reduce overall recidivism rates and provide a net benefit to Washington will depend, in part, on the effectiveness of this resource re-allocation that DOC has made pursuant to the legislative direction of the OAA.

Figure 2
DOC Supervision Hours Budgeted per Offender per Month, by RMI Level

Source: DOC Briefing Document, House Criminal Justice and Corrections Committee, December 6, 2002.

Part Two: Recidivism Rates for the Pre-OAA Comparison Group

As noted, it is too early in the evaluation of the OAA to calculate the recidivism rates of the OAA offenders being followed in this study. We can, however, begin to track the recidivism rates of the offenders in the pre-OAA period who were given the LSI-R. A significant premise of the OAA is that it is possible to use "an objective instrument supported by research" to assess the risk that offenders will re-offend once they are back in the community. DOC selected the LSI-R as its research-based assessment tool and began using it in 1999, a few years before the OAA took effect.
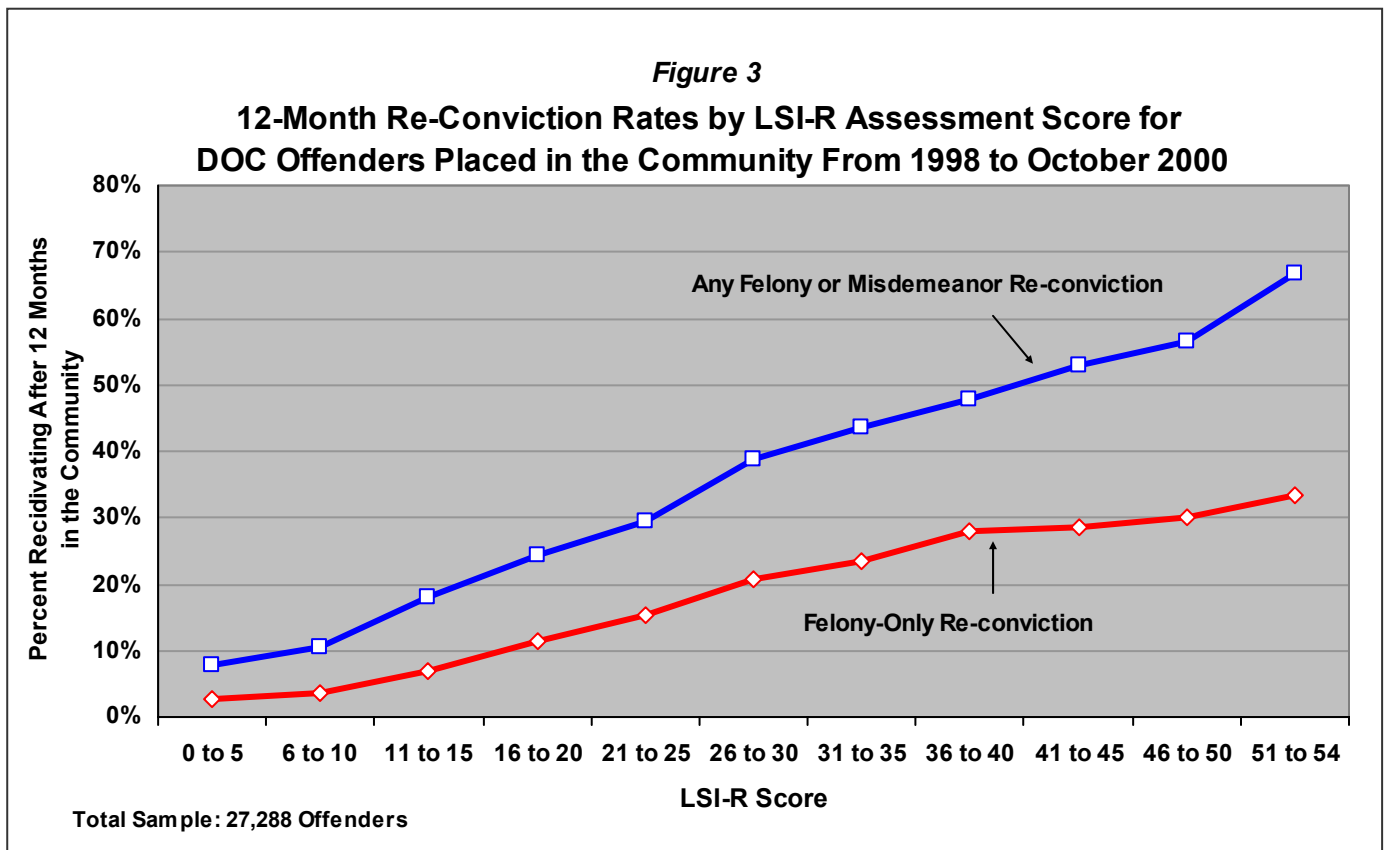
In this report, we present recidivism information on 27,288 pre-OAA offenders who had an LSI-R administered to them during 1999 or 2000. Using Washington's court-based criminal conviction databases, we then followed each of these offenders to see how many were re-convicted for another crime. We tracked this group of offenders for the first 12 months after they were placed in the community. The "community placement" could have occurred as the result of leaving prison or being placed directly in community supervision

following sentencing. We also allowed 12 months for the court to process and adjudicate any subsequent convictions these offenders may have had in Washington.

Figure 3 plots the 12-month recidivism rates for this group of pre-OAA offenders. The figure shows 12-month recidivism rates arranged by the LSI-R score an offender received when DOC assessed them. Two types of recidivism rates are shown on the chart: whether the offender was re-convicted for *any* felony or misdemeanor offense in Washington, or whether the offender was re-convicted only for a felony offense.

As Figure 3 indicates, there is a very strong relationship between an offender's LSI-R score and the chance he or she will be a recidivist within 12 months after being placed in the community. For example, about 11 percent of low-risk offenders with an LSI-R score of 6 to 10 recidivated for any type of felony or misdemeanor in the 12 months following placement in the community. On the other hand, about 53 percent of high-risk offenders with LSI-R scores of 41 to 45 recidivated during the first 12 months on community supervision.

The average LSI-R score for this sample of 27,288 DOC offenders was 21.5. The average

Figure 3

12-Month Re-Conviction Rates by LSI-R Assessment Score for DOC Offenders Placed in the Community From 1998 to October 2000

Total Sample: 27,288 Offenders

12-month recidivism rate for any felony or misdemeanor offense was 28.7 percent, and the average felony-only recidivism rate was 14.4 percent.

Figure 3 shows that there is a clear association between an offender's LSI-R score and the probability that he or she will be re-convicted. This finding should be regarded as preliminary since we will follow these offenders for a longer period of time to assess more fully how well the LSI-R predicts long-term recidivism. Subsequent reports will also analyze these recidivism data thoroughly. The early results shown here, however, can be regarded as a positive preliminary indication that the LSI-R provides a reasonable way to measure an offender's likelihood of re-conviction.

Part Three: Creating the Comparison Group for the OAA Evaluation

In this section, and in the technical appendix, we describe the steps we have taken to date to create a comparison group for use in the evaluation of the OAA. As we report, our initial efforts have not been entirely successful and we outline the next steps we will take, during 2003, to set the stage for the evaluation of OAA outcomes.

After sufficient time has passed, we will be able to evaluate the OAA by comparing the outcomes of the offenders supervised under the Act with a group of offenders in the period prior to OAA implementation. This evaluation strategy is known as a "pre-post" research design. This type of evaluation is not the "gold standard" of research designs. In an ideal research design, offenders would be randomly assigned to the OAA and non-OAA groups, thereby increasing the chance that any observed differences in recidivism outcomes are solely the effect of the OAA. As in most "real world" situations, however, random assignment is not possible for the OAA evaluation since the Act went into effect statewide at one time, rather than in phases.

Despite the fact that random assignment is not possible for the OAA evaluation, the Institute's research design for the Act's evaluation will still be of high quality. As mentioned, DOC began using the LSI-R assessment instrument several years before the OAA went into effect. This fortunate fact will enable the Institute to construct

a reasonable comparison group to use in the evaluation of the OAA.

The Institute uses a 5-point scale to judge the quality of research designs. The purpose of the scale is to measure the degree to which a study's results are reliable. This scale, based on the work of researchers at the University of Maryland,[11] was used by the Joint Legislative Audit and Review Committee in a previous study of the adult corrections research literature.[12] The Institute has also used it to evaluate the "what works" research literature on policies and programs that attempt to reduce crime.[13] On this 5-point scale, a rating of "5" reflects a random assignment evaluation in which the most confidence can be placed. As the evaluation ranking gets lower, less confidence can be placed in any reported differences (or lack of differences) between the program and comparison groups.

Our research strategy for the OAA will produce a "level 3" or "level 4" study on this 5-point scale. This means we can be fairly certain any conclusions we make regarding the effectiveness of the OAA will be valid.

Two-Step Selection Process. Our initial efforts to select the pre-OAA comparison group involved two analytical steps. We summarize the process here, and report the statistical details in the technical appendix to this report, beginning on page 10.

First, we began by examining information in the post-OAA period. We constructed statistical models that predict which OAA offenders are classified by DOC as RMA, RMB, RMC, or RMD. The models were developed with a post-OAA data set that includes actual RMI levels as determined by DOC. Then, using the information on those offenders' LSI-R scores and other criminal history variables, we built multivariate logistic models that

---

[11] L. Sherman, D. Gottfredson, D. MacKenzie, J. Eck, P. Reuter, S. Bushway (1997) *Preventing Crime, What Works, What Doesn't, What's Promising*, Washington: U.S. Department of Justice, Chapter 2.

[12] 1998 Performance Audit report on the Department of Corrections prepared by Washington's Joint Legislative Audit and Review Committee (JLARC). JLARC retained the University of Maryland researchers to judge the overall results and methodological quality of different United States research studies conducted in the adult corrections field.

[13] The scale is described in S. Aos, P. Phipps, R. Barnoski, R. Lieb (2001) *The Comparative Costs and Benefits of Programs to Reduce Crime*, Olympia: Washington State Institute for Public Policy; available at: http://www.wa.gov/wsipp/crime/pdf/costbenefit.pdf.

predict the offender's RMI classification.[14] This OAA data set is described in the Institute's 2002 evaluation report. It includes 9,313 offenders classified by DOC after April 2001.

It is important to note that DOC determines which offenders are classified as RMA, RMB, RMC, and RMD based on a number of factors that we did not include in these OAA multivariate models. As described earlier in this report, DOC assigns these classifications based on an offender's LSI-R score and a number of other factors relating to the harm done by the offender in the past. We did not include these non-LSI-R factors in our models because that information was not available in the pre-OAA period. Thus, the basic requirement for the data that we used in our models of OAA classification was that the same information had to be available (in DOC's electronic databases) in the pre-OAA period.

One significant problem emerged at this stage of the procedure. During the first procedural step, we were unable to develop a good predictive model to estimate which offenders were classified by DOC as being an RMA versus an RMB offender. We were limited in the information we could use to implement DOC's harm-done rules by the restriction, noted earlier, that the same information had to be available electronically in the pre-OAA period. Since much of the distinction between an RMA/RMB classification is made with this OAA information, it is apparently not possible to develop a reliable predictive model without using these data.

Thus, the models that the Institute has developed (to date) predict three categories of DOC offenders: a combined RMA/RMB group, an RMC group, and an RMD group. This limitation should not be too serious for the overall evaluation of the OAA because, as shown on Figure 2, in terms of DOC resources, the RMA and RMB groups are comparable. For example, Figure 2 shows that DOC uses 10.6 hours of community supervision staff time per month per offender for the RMAs and a similar 9.5 hours per month for the RMBs. These figures contrast sharply with 3.6 hours for the RMCs and 0.8 hours for the RMDs.

Thus, unless we are able to develop better models as the evaluation of the OAA progresses,

the main three tests for the evaluation of the OAA will involve comparing recidivism and other outcomes for the combined RMA/RMB group before and after the OAA and the RMC and RMD groups before and after the OAA.

In the second step of the selection process, we took the coefficients from the models developed in the first step and applied them to a group of DOC offenders in the pre-OAA period. The requirements for this pre-OAA data set were as follows:

1. Since the LSI-R is central to the RMI classification system, the fist data requirement for the pre-OAA group was that an offender had to have at least one LSI-R administered and recorded in DOC's electronic files prior to October 1, 2000. October 2000 was chosen because that is when DOC began to classify offenders with its (then) new RMI system; we used this date to separate the pre-OAA period from the OAA period.[15] DOC first began to administer the LSI-R in 1999.

2. From this group, we selected only those offenders with a date of release to the community[16] on or after January 1, 1998. This was done to ensure the bulk of the pre-OAA community supervision occurred relatively close to the OAA period. To be included in the comparison group, the offender also had to be released to the community prior to October 1, 2000 (to separate the pre-OAA period from the OAA period).

3. A third requirement for selection into the pre-OAA comparison group is that an offender could not have been subsequently classified by DOC's RMI system unless the RMI classification was the result of a new conviction or return to prison, in which case that subsequent event would be a recidivism event for the pre-OAA offender. This restriction was used to ensure the pre-OAA group only included offenders who did not receive OAA supervision, treatment, or sanctioning.

---

[14] Our consultant for this project, Northwest Crime and Social Research, Inc., assisted us in this modeling. Northwest Crime and Social Research, Inc., 215 Legion Way SW, Olympia, WA 98501, website: www.nwcsr.com.

[15] Note that for the OAA group, we use RMI dates after April 2001. Some OAA offenders were "RMI-classified" prior to this date but, according to communications with DOC, full training of staff on the use of the RMI system was not completed until April 2001. Thus, for this evaluation, the pre-October 2000 period is defined as the pre-OAA period, while the post-April 2001 is defined as the OAA period.

[16] The "release to community" status can stem from either (a) being released from prison, or (b) being placed directly in community supervision as a result of the sentence.

4. Finally, we excluded from the pre-OAA group those offenders released to the community in locations outside of Washington, since these offenders would not have received the OAA supervision, treatment, or sanctioning.

After applying these four selection rules, we had a pool of 27,323 offenders eligible for inclusion in the pre-OAA comparison group. We then applied the regression coefficients to these offenders' LSI-R and criminal history variables to estimate which of the pre-OAA offenders would have likely been identified as either RMA/RMB, RMC, or RMD prior to implementation of the OAA.

Characteristics of the RMA/RMB, RMC, and RMD Groups

After we selected the pre-OAA comparison group and assigned them an estimated RMA/RMB, RMC, or RMD classification following the procedures described, we compared the groups to see how similar they are based on a number of significant characteristics known to affect recidivism. Ideally, there would be no statistically significant differences between the pre-OAA comparison groups and the OAA groups on any of these characteristics. A lack of significant differences would mean that the groups—*as a whole*—are reasonably comparable and any observed differences in the outcomes among the groups, such as recidivism rates, could be attributed to the effects of the OAA.

Table 1 provides a statistical snapshot of the pre-OAA and OAA groups, arranged by the three

classification levels. Unfortunately, Table 1 reveals that there are significant differences between the estimated pre-OAA group and the actual OAA group. For example, the actual OAA RMA/RMB group has an average LSI-R score of 33.8, while our estimated pre-OAA RMA/RMB comparison groups have an average LSI-R score of 36.4. This is a statistically significant difference. Ideally, there would be no difference between the two groups on this important characteristic. As can be seen in Table 1, the actual OAA RMA/RMB group is generally less risky than the estimated pre-OAA RMA/RMB comparison group. There are also differences between the estimated pre-OAA RMC group and the actual OAA RMC group, although the direction of the differences is split between the pre-OAA RMC group being more and less risky. For the RMD classification, the estimated pre-OAA group is generally more risky than the actual OAA RMD group.

The direction of these differences, as shown on Table 1, is expected. That is, the general net effect of the "harm-done" decision rules in DOC's RMI system is to move lower-risk offenders (where risk is based solely on the LSI-R score) to higher levels of supervision. That is, when we created the pre-OAA comparison group, we did not have the benefit of being able to employ the actual data DOC uses to "override" the LSI-R and re-classify people into different RMI levels. The net effect of these harm-done rules is to move offenders with lower LSI-R scores into higher RMI supervision levels; this effect is confirmed by the results shown in Table 1.

Table 1 Characteristics of RMI-Classified Offenders For the Estimated Pre-OAA Comparison Groups and the Actual OAA Groups						
	RMA/RMB Classified Offenders		RMC Classified Offenders		RMD Classified Offenders	
	Pre-OAA	Actual OAA	Pre-OAA	Actual OAA	Pre-OAA	Actual OAA
Number in Group	2,935	2,450	10,157	4,505	14,231	3,047
Average Age (years)	32.6	32.5	32.3	32.2	31.1	31.8 *
Percent Male	89%	87%	79%	76% *	74%	72% *
Percent European	62%	71% *	72%	78% *	84%	80% *
Percent African	31%	21% *	20%	16% *	10%	13% *
Average LSI-R Score	36.4	33.8 *	26.7	30.0 *	14.8	15.8 *
Prior Adult Felony Convictions	3.7	3.3 *	2.9	3.1 *	2.0	2.1 *
Average Sentence Time for Current and Prior Convictions (months)	30.5	27.8 *	14.6	12.6 *	6.0	8.2 *

Note: An asterisk (*) indicates a statistically significant difference with a p-value less than .05 from either a t-test or a chi-square test.

These differences in the pre-existing composition of the comparison and OAA groups means that—without further adjustment—erroneous conclusions could be drawn from the evaluation of OAA outcomes by simply comparing the unadjusted results of the two groups. For example, since the actual RMA/RMB OAA group is slightly less risky than the estimated pre-OAA RMA/RMB comparison group, we would expect unadjusted recidivism rates to be, on average, somewhat lower for the OAA group, regardless of whether the OAA has any effect or not.

There are three possible statistical solutions to avoid drawing these erroneous conclusions. First, since there are pre-existing differences among the groups, one option is just to use multivariate statistical models to adjust for the differences in the composition of the groups. This is the standard methodological approach when there are pre-existing differences between groups. The difficulty with this solution is that the multivariate controls may not be sufficient to estimate the observed and, perhaps, unobserved selection factors that DOC assigns in the RMI classification process. Still, by using statistically-adjusted results, the research design would generate a "level 3" research quality measure on the 5-point research design scale mentioned earlier.

Second, during 2003, the Institute will investigate the feasibility of using alternative methods to construct the comparison group. In particular, we will employ specific matching algorithms to create a comparison group. These matching procedures should eliminate some of the differences shown on Table 1. These methods select a comparison group by performing a case-by-case statistical match to identify an individual pre-OAA offender for each OAA offender based on a number of key characteristics of interest.[17]

Our third research strategy will take a different approach. The models we have built to date do not provide an adequate comparison group because they cannot utilize all the information DOC uses in its RMI system. In particular, DOC employs a number of harm-done and other decision-making rules, in addition to the LSI-R, to decide whether an offender should be classified as an RMA, RMB, RMC, or RMD offender. Because this supplemental information is not available for the pre-OAA period, the models we

have built to select the pre-OAA comparison groups are limited in their ability to provide reliable predictions for constructing the comparison group.

If, however, these other DOC "harm-done" decision-making rules are not related to the probability of re-offending, but are in fact related to how DOC classifies offenders, then it is possible to employ the statistically powerful "instrumental-variables" methodology to evaluate the OAA.[18] That is, if substantial groups of offenders are being supervised at different RMI levels for reasons not related to the probability of re-offending (namely, the harm-done rules related to prior, not future, behavior), then a strong "level 4" evaluation design using instrumental variables can be used to isolate the effect of different levels of supervision and treatment on the probability of recidivism. This statistical approach obtains consistent estimators because it takes advantage of the variation in the OAA's main policy variable (the resource allocations associated with RMI classification levels) caused by exogenous factors (DOC's harm-done criteria) that do not otherwise relate to the outcome variable of interest (recidivism rates). During 2003, this instrumental variables approach will be developed fully for the OAA evaluation.

---

[17] Matching methods are covered in P. Allison (1999) *Logistic Regression Using the SAS® System: Theory and Application*, Cary, NC: SAS Institute Inc., Chapter 8.

[18] Standard references include: J. M. Wooldridge (2000) *Introductory Econometrics, A Modern Approach,* Cincinnati: South-Western College Publishing; J. M. Wooldridge (2002) *Econometric Analysis of Cross Section and Panel Data,* Cambridge: The MIT Press; and W. H. Greene (2000) *Econometric Analysis, 4th Ed.,* New Jersey: Prentice Hall.

As described in the main report, we adopted a modeling approach to estimate which offenders in the pre-OAA period would have been classified as RMA/RMB, RMC, or RMD offenders. This appendix describes the model developed (to date) to estimate these categories.[19]

The process of choosing a comparison group involved two steps.

- First, we statistically modeled the DOC assignment process for OAA offenders. Since we know an offender's actual RMI classification, we used available data to see what factors were associated with that classification decision. Several statistical models were developed, and RMIs predicted by the statistical models were then compared with actual RMIs. The model that most accurately predicted actual RMIs was selected to be applied to the pre-OAA group.

- Second, we applied the results of that statistical model to pre-OAA offenders, and predicted what RMI they would have been assigned, had the OAA been in effect in the pre-OAA period.

Step 1: Statistical Modeling for OAA Offenders

Modeling Technique. The cumulative logistic regression model was chosen as the statistical technique to model the DOC assignment process for several reasons. First, logistic rather than ordinary least squares regression was chosen because the dependent variable is categorical rather than continuous. We modeled whether an offender was in one of three categories, 1) RMA or RMB, 2) RMC, or 3) RMD.

Second, the cumulative logistic model is designed for situations where the dependent variable has more than two categories and there is an inherent

ordering in these categories. In this case, the risk management categories assigned by DOC do have an inherent ordering: RMAs and RMBs are believed to require more supervision than RMCs, while RMDs require the least supervision. To ignore that ordering would lead to problems in both the estimation of parameters and the interpretation of results.

We estimated models using SAS software. SAS's PROC LOGISTIC automatically estimates a cumulative logistic model anytime there are more than two categories on the dependent variable.

Model Specification: The Dependent Variable. The objective of the multivariate models is to predict an offender's DOC-assigned RMI level. In determining an offender's RMI, DOC uses several pieces of information. The first is an offender's LSI-R score, a standardized instrument that is designed to assess an offender's risk to recidivate. In addition to the LSI-R, however, DOC uses a series of other factors, primarily related to the "harm done" of the offenses committed by the offender.[20] These harm-done measures do not always correspond with an offender's LSI-R score. For example, an offender might have a low LSI-R score yet have committed a violent act against a vulnerable person. That offender would have been placed in either the RMA or RMB group—groups that receive high levels of supervision under the OAA.

The problem the DOC method of classification poses for the evaluation design is that the harm-done factors are not available electronically, if at all, for the pre-OAA offenders. Thus, they cannot be included in our pre-OAA statistical models. Because, in some cases, these harm-done factors override the influence of other factors, such as the LSI-R score, our efforts at prediction are not perfect, and RMAs cannot always be distinguished from RMBs, given the information

---

[19] Most of this technical appendix was prepared by the Institute's consultant on this project, Northwest Crime and Social Research, Inc., 215 Legion Way SW, Olympia, WA 98501, website: www.nwcsr.com.

[20] These "harm-done" factors include, for example, whether an offender 1) is a level 3 sex offender, 2) is a dangerously mentally ill offender, 3) has committed violence toward strangers, vulnerable people, institutions or groups, and 4) is a demonstrated threat to victims.

that we can use. Therefore, we combine the RMAs and RMBs into one category of offenders for the purpose of the evaluation of the OAA.

Model Specification: The Independent Variables Used. In the multivariate models, we use an offender's LSI-R score (grouped into the following four categories, 0–23, 24–31, 32–41, greater than 41); an individual LSI item indicating whether an offender has an official record of violence; Institute measures of criminal history (number and type of felonies and misdemeanors prior to LSI-R administration), and demographic data (age, race, sex).

Modeling Results. Table A-1 shows the results of the model that best fits the data on OAA offenders. The easiest way to interpret the results of a cumulative logit model is to focus on the odds ratios. The odds ratio for any independent variable can be interpreted as the effect of that variable on the odds of being in a higher rather than a lower category. In this case, the higher categories indicate RMIs that involve more, rather than less, DOC supervision. So, our results show that a one-unit increase in LSI-R score (recall that LSI-R score was divided into 4 categories 0–23, 24–31, 32–41, or greater than 41) and having a prior sex offense, greatly increases the odds of being in a higher category and of receiving extensive supervision.

Modeling this process in SAS also allows us to assign a predicted RMI for OAA offenders. An option in PROC LOGISTIC produces the predicted probabilities of being in each of the RMI categories for each observation in the dataset. The category with the highest predicted probability is chosen as the predicted value of RMI for each offender.

Table A-2 shows a comparison of predicted RMIs with actual RMIs, for OAA offenders. There are several findings to note: first, of the 2,400 actual RMAs or RMBs, 1,520, or 63 percent, were predicted to be RMAs or RMBs. Of the 1,989 predicted RMAs or RMBs, 1,520, or 76 percent, were actually an RMA or RMB. The accuracy of the predictions for the other categories was even higher.
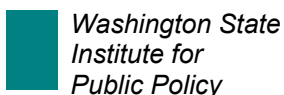
Step 2: Application of the Model to Pre-OAA Offenders

Identifying Pre-OAA Offenders. After the model was developed, the estimated coefficients could be applied to an OAA group of offenders. The selection methods we employ for this process are described fully on page 7 in this report.

Calculating Predicted Probabilities and Assigning a Predicted RML. Observations for pre-OAA offenders are included in the dataset when modeling OAA RMIs. Since pre-OAA offenders have no value for the dependent variable, they are not included in the estimation process. However, PROC LOGISTIC calculates predicted probabilities of being in each category of the dependent variable for all observations, regardless of whether the dependent variable is missing. Thus, pre-OAA offenders are given predicted probabilities of being in the various RMI categories. The category with the highest predicted probability is selected as the predicted RMI.

Table A-1
Results of Cumulative Logistic Regression Model
Predicting Risk Management Level for OAA

Independent Variable	Parameter Estimate	Std Error	Chi-Sqr	P-Value	Odds Ratio
Intercept (RMA/RMB)	-9.08	0.17	2857.8	0.0001	
Intercept (RMC)	-4.55	0.13	1226.4	0.0001	
Age 18–30 (compared with age > 45)	-0.07	0.08	0.78	0.38	0.93
Age 31–45 (compared with age > 45)	-0.12	0.08	2.26	0.13	0.88
Male	0.17	0.06	7.09	0.008	1.19
Asian (compared with Black)	-0.08	0.17	0.22	0.63	0.92
White (compared with Black)	-0.29	0.07	17.2	0.0001	0.75
Other Race (compared with Black)	-0.27	0.14	3.8	0.05	0.77
LSI-R Score	2.77	0.05	3724.5	0.0001	15.9
LSI-R Official Record of Violence	0.75	0.06	158.33	0.0001	2.11
Prior Person Offenses (Y/N)	1.23	0.06	388.57	0.0001	3.43
Prior Sex Offenses (Y/N)	3.49	0.14	633.65	0.0001	32.83
Prior Homicide	1.84	0.32	32.28	0.0001	6.28
Model Fitting					
Percent Concordant Pairs	91.8%				
Max-Rescaled R-Sqr	0.73				
N=9,313					

Table A-2
Comparing Predicted and Actual Risk Management Levels
for OAA Offenders

	Actual Risk Management Level									
	RMA or RMB			RMC			RMD			
Predicted Risk Mgmt Level	N	Row%	Col%	N	Row%	Col%	N	Row%	Col%	TOTAL
RMA or RMB	1,520	76%	63%	468	24%	11%	1	0.05%	0.04%	1,989
RMC	732	16%	31%	3,593	80%	87%	159	4%	6%	4,484
RMD	148	5%	6%	55	2%	1%	2,637	93%	95%	2,840
TOTAL	2,400		100%	4,116		100%	2,797		100%	9,313