Volume 70 Number 2

Federal PROBATION
a journal of correctional philosophy and practice

Home

From Theory To Practice: The Lifecycle Document For The Results-Based Management Framework For The Federal Probation And Pretrial Services System

Endnotes

[The federal probation and pretrial services system is making steady progress towards a comprehensive outcome-measurement system [1] *that will allow the federal judiciary's policymaking body, the Judicial Conference of the United States, and chief probation and pretrial services officers and system managers to make decisions based on ongoing empirical analysis of reliable and pertinent data. We expect to have much of the data collection infrastructure in place in fiscal year 2007 and all of it in place in fiscal year 2008. We will then be able to identify various recidivism-reduction strategies used by different districts and will be able to track large cohorts of offenders to determine what strategies appeared to make a difference under what circumstances for what types of offenders.*

Pending completion of their own internal outcome-based system, federal probation and pretrial services officers are eagerly adopting research-supported recidivism-reduction strategies identified in the growing body of research into state and local programs known as Evidence-based Practices. [2] *When the federal outcome-based system is fully implemented, it will be possible to determine whether those evidence-based practices made a significant difference in measurable outcomes and, if so, whether they should be implemented in other districts.*

What follows is a report prepared by Caliber, An ICF Consulting Company under contract with the Administrative Office of the United States Courts. This report provides some useful background and describes precisely how we have built and will continue to build our outcome-based system. Much of the body of the report is reproduced below, except for most of the attachments and section 6, which details nuts-and-bolts implementation.]

1. Project Background

The federal probation and pretrial services system is developing a results-based management framework that will, in the future, allow it to better assess performance—and make programming and resourcing decisions—based on what it accomplishes rather than solely on what it does. The flow chart below shows the steps involved in developing the framework, and highlights where we are in the process.

This focus on results, and the work done to date to define the system's mission, goals and desired outcomes, stem from a number of complementary influences and projects.

1. In 1999, the Administrative Office entered into a contract with a team of independent consultants, led by IBM, to conduct a strategic assessment of the federal probation and pretrial services system. The overarching recommendation from that assessment — presented first to the Administrative Office in 2003—was that the federal probation and pretrial services system become a results-driven organization with a comprehensive

performance measurement system.

2. In 2000, the AO Director appointed an Ad Hoc Supervision Work Group comprised of supervisors, deputies and chiefs from seven districts and a representative of the Federal Judicial Center to update the supervision policy monographs. As part of its work, the group reviewed relevant statutes and mission statements to identify the desired outcomes and goals to be served by the pretrial services and post-conviction supervision functions. These outcomes and goals were incorporated into revised supervision policy documents approved by the Judicial Conference of the United States in 2003.

3. Strategic planning sessions were conducted at the 2000 and 2002 Federal Judicial Center's National Chiefs Conferences. The 2000 conference produced a "Desired Futures" roadmap the first element of which was: "Desired Outcomes are clear, measured and results are communicated." The 2002 conference resulted in a "Charter for Excellence" that sets forth broad system goals and values (see Attachment 1).

4. In September 2003, one of the IBM strategic assessment consultants facilitated a strategic planning session at a meeting of the Chiefs Advisory Group to translate the broad "Charter for Excellence" statements into more specific "Operational Goals."

The operational goals developed by the Chiefs Advisory Group were combined with the desired outcomes set forth in the revised supervision monographs to form the basic structure of the results-based management framework (see Attachment 2). This concluded the initial goal-setting stage of the framework development process. [3]

The current stage of the process is technical: The development of operational definitions and associated measures for each "desired outcome;" and of statistical approaches to analyze the information that will assure "apples-to-apples" comparisons and allow benchmarking with other programs. To assist in this phase:

- In March 2004, the AO director appointed an Ad Hoc Expert Panel on Outcome Measurement Methodology. This panel is comprised of the directors of research for the Federal Bureau of Prisons, the Federal Judicial Center and the District of Columbia's Court Services and Offender Supervision Agency; and academics and Administrative Office staff with extensive backgrounds in criminal justice/substance abuse research and government performance measurement systems. The panel met twice in 2004 and produced recommendations for measuring the concepts of recidivism and substance abuse (see Attachment 3).

- In April 2005, the AO entered into a contract with Caliber Associates to provide additional support in developing, coordinating and presenting recommendations for the technical aspects of the results-based framework.

This final product from the technical phase includes recommendations, to be circulated to system staff and stakeholders for broad system comment, that span the lifecycle of the results-based management framework. The report addresses:

- How to measure a variety of outcomes—including offender compliance, positive change, and crime reduction;
- What data are needed to construct the recommended measures;
- What analytical methodologies can be used to assess how these results are affected by supervision interventions as well as a variety of case, offender and community factors?
- What tasks are necessary to fully implement the framework; and
- How to institutionalize the framework within the federal probation and pretrial services system?

The recommendations represent "state of the art" measurement and analytical approaches that are being used by other performance-based systems, program evaluations and/or academic research in criminal justice and related areas such as substance abuse.

2. Post-conviction Supervision Logic Model

Building on the results of the goal-setting stage of this project, the next step was to develop a logic model that depicts the underlying assumptions about how "what the system *does*" affects what it is trying to *accomplish*; and what other factors—e.g., characteristics of the offenders to be supervised, the requirements and restrictions of their sentences, and the system resources devoted to carrying out the supervision mission—are expected to influence this relationship. The logic model for post-conviction supervision is used throughout this document to illustrate the technical concepts that are incorporated into the framework design.

This logic model has been refined twice since its development following the goal setting stage based on input from the Expert Panel and Caliber Associates. It will continue to be a work in progress that evolves to incorporate feedback from system staff and stakeholders, and results from empirical testing of the posited relationships. A similar logic model was developed for pretrial services; which will be incorporated into the framework and follow a similar evolution process.

2.1 Components of the Logic Model

The post-conviction supervision logic model has six components: inputs, process (activities), process outcomes, intermediate outcomes, ultimate outcomes, and mission. Each component is described below.

Inputs

Inputs are characteristics of the offender population and the working environment that are hypothesized to affect expected outcomes *regardless* of system interventions. For example, prior research indicates that offenders with a lengthy prior record are more likely to become re-involved in criminal activity than those with no or a minimal prior record. This leads to a working assumption that, regardless of supervision interventions, districts that have a high percentage of first offenders will have a lower recidivism rate than those with a low percentage of first offenders.

Inputs are used in the analytical model as "control" variables to account for the effects of factors that explain differences in outcomes across offices, districts and time that are not related to system interventions. They may also be used as stratification categories to display outcomes based on key groupings, e.g., by type of supervision (probation/parole/supervised release) in reports.

The current model includes as inputs those factors identified in the research and program evaluation literature as related to criminal justice goals. These include:

- Offender characteristics (e.g., prior record, employment; family/community connections, demographics);
- Characteristics of the instant offense (e.g., class and category);
- Sentence parameters (e.g., length of prison and supervision terms imposed and served, conditions imposed);
- Office/community characteristics (e.g., location, size, socio-economic indicators);
- Officer characteristics (e.g., experience, demographics, education);
- Supervision resources (e.g., supervision staffing, contract budgets, technological support).

The inputs categories will be further defined and the categories and their specific elements assessed for adequacy by system staff and stakeholders as part of the framework implementation. These inputs will also be applied, as appropriate, to the pretrial services model.

Process (Activities)

Process refers to activities undertaken by the system—practices, programs and interventions—that implement the supervision function. As an example: An officer conducts an initial assessment investigation, identifies lack of stable employment as a risk, and refers the offender for job counseling or to a job referral agency. In the analytical model, the process variables define "what we do" for purposes of assessing the basic relationship of how "what we do," relates to what we are trying to accomplish.

The current logic model includes only the most general process categories, e.g., investigation, assessment, monitoring, referral, and assistance. Detailed input on the specific processes that should be included in the model will be sought from system staff and stakeholders—the experts in identifying and defining salient system activities—as part of the initial next step in implementing the framework.

Process Outcomes

Process outcomes describe *offender* actions that occur as a result of system activities. For example, in response to an employment referral, the offender registers with an employment service or completes "x" hours of employment counseling.

Process outcomes enter the analytical model as both an outcome of the service delivery process and as an input (control) for assessing interim and ultimate outcomes. For example, "number of hours of employment counseling" is a measure of how successful an officer's employment referrals are in engaging offenders in employment services. This measure is also a "control" when addressing the extent to which an interim outcome, such as improved employment, might be attributed to the supervision intervention.

The model includes only broad categories of compliance with each of the four major types of conditions: Restrictions, correctional programming, service, and financial. These will be expanded as a result of comments received from system staff and stakeholders.

Intermediate Outcomes

Intermediate outcomes are changes in offender behavior that are themselves desirable and believed also to be precursors of the ultimate outcomes. Specifically, the intermediate outcomes in the model are defined as a desirable *change* in a circumstance that has an *empirically proven relationship* to successful supervision and that is within an officer's *authority and sphere of influence*.

Three intermediate outcomes are included in the current model: Reductions in substance abuse, improvements in employment, and improvements in other life skills. Each of these changes is believed to relate to the likelihood of criminality during the period of supervision and beyond. Each is also expected to relate to certain "sentence execution" outcomes (e.g., improved employment, enhanced earning capacity, more money to pay restitution).

The list of intermediate outcomes currently in the model does not include all of the operational goals that emerged from the Chiefs Advisory Group's strategic planning session in 2003. This results from the limitations that were subsequently attached to the definition of intermediate outcomes. For example, "Improvements in Mental Health" is *not* listed as an intermediate outcome because there is no empirical association between general mental health problems and criminality. This does not mean that mental health issues will not be considered, but rather that their relationship to ultimate outcomes will be considered in terms of offender and sentence characteristics (e.g., mental health needs and conditions as inputs), referrals for mental health counseling (process), and the offender's participation in that programming (process outcomes).

Ultimate Outcomes and Mission

The ultimate outcomes are set forth in *The Supervision of Federal Offenders, Monograph 109*, which establishes Judicial Conference policies related to post-conviction supervision. These outcomes are: To execute the sentence and to protect the community during the period of supervision and beyond.

These outcomes stem directly from the system mission endorsed by the Judicial Conference in September 1993—To protect the public and to assist in the fair administration of justice—supplemented by the statutory provisions that establish the duties of probation officers and the purposes that community sentences are to serve.

Two of the ultimate outcomes—minimizing criminal activity during the period of supervision and beyond—relate to the system mission to protect the public. The other ultimate outcomes measure the impact of compliance with release conditions (e.g., restoration of victims). These serve as direct measures of sentence execution and surrogates for the mission to assist in the fair administration of justice—with no expectation that they will affect the system's public protection mission. Examples from the literature of more precise operational measures for process outcomes, intermediate outcomes, and ultimate outcomes are provided in Section 3 and further delineated in Attachment 5: Key Element Definitions and Attachment 6: Data Matrix.

2.2 Relationships among Components

The arrows in the logic model indicate the specific expected relationships between components that the analytical model will be designed to test. As described in Section 4, statistical techniques will be applied to test the relationships depicted.

The analysis will test a complete thread of the model, starting from left to right. Basic and advanced techniques will be used to test both direct and indirect and unidirectional and bidirectional relationships, while controlling for inputs that are primarily static and outside the control of the officer. The results will move the system beyond a description of the offender population and individual outcomes to a more complex assessment of the "theory of change" and the interconnectedness of process and outcomes for post-conviction supervision. Similar relationships will be tested for pretrial services.

3. Operationalizing Post-conviction Supervision Outcomes

This section further defines the process, intermediate, and ultimate outcomes in measurable terms. In order to empirically test the hypothesized relationships between post-conviction processes (activities) and outcomes of the offender population, it is necessary to first identify appropriate measures for each outcome. A data matrix was developed based on a review of current evaluation research. [4] The matrix organizes each outcome into the following categories:

- Type—The concept of interest.
- Definition—A brief description of each concept. The descriptions are important to ensure standardization in how data are defined across districts.
- How Operationalized—How the concept will be measured.
- Data Element—The specific piece of data to be captured so that certain tests can be performed to help answer questions of interest. Data may be reported in days, weeks, dollars, cents, or by selecting "yes" or "no" options or other response categories.
- Level of Measurement—The level at which the data will be measured (nominal, interval, ratio, ordinal). This is important for determining the type of analysis that can be performed using each measure.
- Data Source—Where the identified data may be obtained. In some cases, the data may not currently be collected and therefore, the data source will need to be determined by system staff and stakeholders.

3.1 Process Outcomes

A process outcome represents the immediate outcome for the offender as a result of system activities. The four process outcomes are: compliance with restrictive conditions, participation in correctional programming, compliance with service conditions and compliance with financial conditions. Examples of the types of data that could be collected for each outcome are described below.

- Comply with restrictive conditions (e.g., home confinement conditions, halfway house placement, employment conditions, prohibition of contact with victim/minor/associates, remote location monitoring, and nighttime/weekend jail requirements). The process outcome measures for all conditions will include: a dichotomous measure of compliance (complied or did not comply) and number of noncompliant events involving condition. An additional measure for any condition with an associated time component will be calculated based on days imposed vs. days completed. The process outcomes will be calculated separately for each restrictive condition and for the restrictive category as a whole.

- Participate in correctional programming—Measures of this process outcome will include: number of days from start to end of program (duration); number of hours of service (per week and total); number of sessions attended vs. sessions scheduled; end-of-treatment provider assessment of quality of participation (5-point scale); and completion status (successful/unsuccessful completion). The results will be presented by type of program (substance abuse, education/employment/job training, mental health treatment, sex offender, life skills, basics), funding source (no cost, contracted, other government program, or self-insured), and modality of treatment (inpatient, individual, group, family-individual, or family-group).

- Comply with service conditions—Service conditions consist of a requirement for the offender to complete hours of non-paid community service. Compliance with service conditions will be determined by number of community service hours completed vs. the number of hours imposed. A dichotomous compliance status measure (complied or did not comply) can then be calculated based on whether the offender successfully completed the imposed hours of non-paid community service.

- Comply with financial conditions—(e.g., fine, restitution, special assessment, no new debt/credit, cooperate with IRS; child support enforcement) Measures for all conditions will include: a dichotomous measure of compliance (complied or did not comply) and number of noncompliant events involving condition. An additional measure for any financial condition with an associated amount will be: amount expected by payment schedule vs. amount paid. The process outcomes will be calculated separately for each financial condition and for the financial category as a whole.

The process outcomes data described above will be used in the analysis described in Section 4 as both dependent variables (predicted outcome of system activities) and independent variables (predictors of intermediate outcomes).

3.2 Intermediate Outcomes

An intermediate outcome represents the expected immediate result of the process outcomes. Examples of the data to be collected for the three intermediate outcomes depicted in the post-conviction supervision logic model—reduce substance abuse, improve employment, and improve other life skills—are described in greater detail below.

- Reduce substance abuse—Measures of substance use *during the period of supervision* will include drug test results, self-admissions, and substance-related re-arrests. These measures will be used to create a dichotomous measure of substance use (used or did not use). The outcomes will be presented by time of event (before treatment, during treatment and after treatment) and type of substance. The analysis will assess “change” by comparing substance use during supervision with the offender’s status at the start of supervision based on such factors as: prior diagnosis of addiction/abuse (Y/N); evidence of use at time of instant offense based on admission, positive pretrial/presentence drug test, offense involving drug/alcohol use? (Y/N); drug(s) of choice; and number of prior treatment experiences.

- Improve employment—Measures of employment may include: change in employment status (unemployed, unemployed but seeking employment, part-time employment, full-time employment), length of employment (calculated as the percent of time offender was employed during period of supervision), and amount of wages.

- Improve other life skills—There is still a question as to the type of life skills that should be identified as outcomes in the logic model. Specifically, which life skills meet the intermediate outcome criteria: "A desirable change in a circumstance that has an empirically proven relationship to successful supervision and that is within an officer's authority and sphere of influence." Potential straightforward areas are level of educational and new vocational/advocational skills. Can or should the area be expanded to encompass such topics as family stability and community stability without overstepping appropriate bounds on officer authority?

As with process outcomes, the above intermediate outcomes will be tested as dependent variables (the result of participation in correctional programming) and independent variables (predictors of future criminal activity and victim restoration). As shown in the logic model, the relationship (as either independent or dependent variables) between intermediate outcomes also will be tested. Additionally, the intermediate outcomes will be tested as mediating variables between process outcomes (participation in correctional programming) and ultimate outcomes (minimized criminal activity).

3.3 Ultimate Outcomes

An ultimate outcome is the long-term result of the system activities for the offender. The ultimate outcomes also reflect achievement of the mission of the federal probation and pretrial services system. The five ultimate outcomes that best reflect the mission include: minimize criminal activity during the period of supervision, minimize criminal activity beyond the period of supervision, maximize victim restoration, defray costs to the government, and maximize compliance with release conditions. The analysis of data on these ultimate outcomes will help system staff and stakeholders better assess if the missions of protecting the public and assisting in the fair administration of justice are being achieved. Each ultimate outcome is discussed below.

- Minimize criminal activity during the period of supervision—The primary measure of criminal activity during the period of supervision is whether an offender was arrested for a new offense. Technical violations are not counted as a new offense. The analysis will also examine the time to arrest (length of time before the arrest for a new offense). The results will be presented overall and by offense type (e.g., violent, property, drug, public order, weapon, immigration) and offense level (felony, misdemeanor, petty).

- Minimize criminal activity beyond the period of supervision—Similar measures for criminal activity beyond the period of supervision will be reported as described above.

- Maximize community restoration—Community restoration will be measured as the amount of payments made to victim special assessments and, where required, the Victims' Crime Fund. In addition, community restoration will be measured by the amount of fines paid, etc.

- Maximize compliance with release conditions—Compliance with release conditions will include: a dichotomous measure of any noncompliance (Y/N), number of instances of noncompliance, time to first noncompliance (months), and time free of noncompliance at inactivation/termination of supervision (months).

Ultimate outcome data enable system staff and stakeholders to test whether the system activities (processes) are leading to the long-term outcomes that the federal probation and pretrial services system is tasked with achieving. Furthermore, these data will allow system staff and stakeholders to assess how well they are doing at meeting their mission to protect the public and fairly administer justice.

4. Data Analysis Plan

The data analysis plan describes a recommended approach to testing the relationships depicted in the logic model. That is, the plan presents the statistical techniques, progressing sequentially from simpler to more sophisticated levels of analysis, that will provide system staff and stakeholders with both a description of the offender population and outcomes and a more complex assessment of the hypothesized "theory of change" and interconnectedness of process and outcomes (e.g., direct, indirect, unidirectional and bidirectional relationships) depicted in the post-conviction supervision logic model. Specifically, the plan describes basic and advanced statistical techniques that can be applied to test these relationships, while controlling for inputs that are primarily static and outside the control of the probation officer. The analysis plan is organized into three stages: data quality, data reduction, and data analysis. Each is presented below.

4.1 Data Quality

Before any data analysis is conducted, all data will need to undergo standard checks for quality to make sure there are no data entry or transmission errors. Checking for data quality is typically a two-step process that involves detection and then correction of errors in a data set. Cleaning and preparing data is an often neglected but extremely important step in the analysis process. The saying "garbage-in-garbage-out" is particularly applicable where large data sets collected via some automatic methods (e.g., via National PACTS Reporting, National Crime Information Center (NCIC), etc.) serve as the primary input into the analysis. The most common sources of error include data entry errors, such as typing errors, column shift (data for one column being entered under the adjacent column), which often results in invalid responses, general coding errors, which may occur during data collection or entry and may be difficult to detect unless you look for outliers or unusual relationships between variables, and not recoding missing data, which can result in inflated mean scores and the like. The first step to ensuring quality data begins with ensuring clear understanding and procedures for collecting and entering data and systematic review before data are submitted or made available for extraction. Once data are submitted, the next step in the data quality assurance process is to detect and clean the data for errors.

Error Detection

There are three common procedures for detecting data errors that should be followed. These include:

- Review descriptive statistics. Using software, such as SPSS, the following review of descriptive statistics can identify data errors:
    - Look at minimum and maximum values to determine if data fall outside the acceptable range.
    - Look for presence of 0's and 999's (or 9999, etc.) shown in frequency tables, graphs or histograms to indicate missing values.
    - Look at means, medians, and standard deviations. For example, if the median differs much from the mean value, it is important to investigate the overall distribution of values for outliers.
    - Assess frequencies. By examining frequencies, it is possible to detect unequal distribution in categories such as age, sex, and race that are outside what would normally be expected for a particular population.
- Conduct logic checks. Errors in data can be detected simply by determining whether the

responses seem logical. For example, you would expect to see 100% of responses, not 110%.

- Examine bivariate outliers. Some data errors only appear when two variables are compared. To detect such error, it is important to look for outliers, or values of a variable that are far different from the expected values. These errors can be detected by examining bivariate associations and scatter-plot graphs to check for deviations in expected associations between variables.

Once the data errors are detected, there are several techniques that should be followed for correcting the errors. These are described below.

Error Correction

Once errors are detected, it is important to know how to handle them appropriately so the data can be analyzed without losing their integrity or robustness. There are slightly different ways to deal with error in independent (or predictor and control) variables and dependent (outcome) variables.

Independent Variables

When there are a minimal number of errors, the values are generally recoded to "missing." What this means is that the suspicious values are counted as missing data since they are not within an acceptable range. If there are many error values, then it is important to check to see if some of the values of the outcomes are the same for missing and nonmissing values for the independent variables. If so, then there is less chance of bias in the analysis. If not, then it is possible that the data is not good and that the variable should be discarded or used with caution. Various imputation-based procedures to fill in missing values (series mean, mean or median of nearby points, linear interpolation, linear trend at point) will need to be considered. Other more complex imputation-based procedures (regression imputation, non-ignorable missing-data models, Heckman's two-step statistical process) may also need to be used. [5]

Dependent variables

If there are few data errors, values can be set to "missing" using one of the imputation-based procedures determined to be most appropriate. However, it is important to use caution when setting many values to "missing," especially if multiple variable analysis will be conducted . It may be necessary to set the error values for the outcome or independent variable to the data set mean or the group mean (maybe by age, type of offender, etc.). This should result in a histogram with a more normal distribution of values. Once quality of the data has been checked and the necessary steps taken to correct for problems or errors, data analysis can proceed from basic to more advanced techniques described below.

4.2 Data Reduction

Data reduction is a process often applied where the goal is to aggregate or amalgamate information contained in large data sets into more manageable and reliable information. Data reduction techniques can include simple tabulations, aggregations, or more sophisticated techniques, such as clustering, principal components analysis, and path analysis. Each of these recommended data reduction methods is described below.

Aggregation

Aggregating or transformations of data are techniques often used to reduce or optimize your data. Data can be aggregated, for example, by subgroups to move from individual case records for thousands of individual offenders to mean scores for subgroups of offender populations based on certain criteria (e.g., criminal history, district, gender, etc.). Additionally data can be transformed by creating dichotomous variables (presence/absence of new offense) from continuous variables (number of new offenses) or composite scores or constructs from multiple measures (risk assessment score). When transforming data, especially creating composite measures or constructs, it is important to use other techniques, such as those described below, to determine which variables should be combined to create a new variable.

Cluster Analysis

Cluster analysis is a multivariate analysis technique that seeks to organize information about variables so that relatively homogeneous or similar groups, or "clusters," can be formed. To use this technique, it is important that the clusters formed be highly internally homogenous (members are similar to one another) and highly externally heterogeneous (members are *not* like members of other clusters). Cluster analysis can be used to combine "similarity" measures as well as measures that are proxies or associations. However, it is first necessary to standardize your data since clustering often involves combining items measured on different scales.

Principal Component Analysis

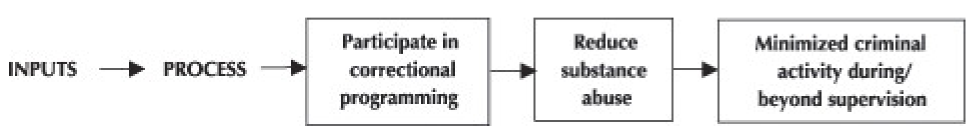
Principal component analysis is used to reduce the number of variables or factors for inclusion in your analysis. Specifically, this method of analysis is used to combine two or more *correlated* variables into a single factor. Principal component analysis helps reduce redundancy in your measures by identifying and combining those that are highly correlated into a new variable, while minimizing the variance around the new variable. [6]

Path Analysis

Once you have arrived at a set of measures that represent the variables of interest (inputs, processes, process outcomes, intermediate outcomes, and ultimate outcomes), it is necessary to determine whether the relationships or paths among those variables presented in the post-conviction supervision logic model can be supported by the data. That is, do the data fit the model? Path analysis calculates a path coefficient, which shows the direct effect of an independent variable on a dependent variable in the path model. This information is then used to calculate a goodness-of-fit statistic. The statistic determines how well the data fit the model, or stated another way, the statistics help identify the best fitting models for the data. [7] For example, based on the analysis, it may be determined that certain predicted or hypothesized relationships shown in the logic model are not supported by the data. Or, the analysis may uncover other relationships (direct or indirect) that are not currently represented in the model. It is important to test these relationships using path analysis or other more sophisticated techniques, such as structural equation modeling (a variation of path analysis involving multiple indicators of variables in the model), to ensure the "big picture" is accurate before attempting to further describe your data and examining more minute relationships among variables (see basic and advanced analysis sections below). [8] Based on the results, it may be necessary to modify the model (add and/or remove arrows depicting relationships, add/or remove variables, etc.).

4.3 Data Analysis

The data analysis stage is presented in two parts: basic analysis and advanced analysis Each is presented below. To help illustrate the rationale behind each recommended analytic technique, examples drawn from the substance abuse "thread" of the logic model shown below are discussed.



Basic Analysis

Once the steps necessary to reduce or optimize the available data have been completed, the first or basic level of analysis to be conducted of the input, process, and outcome data will use descriptive statistics. Descriptive statistics should be used to:

- Describe the basic features of the data
- Provide simple quantitative summaries about the measures
- Provide the basis for subsequent data analysis.

In general, descriptive statistics will describe "what is," or what the data show with respect to a given variable.

Univariate Analysis

Univariate analysis of the data is used to examine across cases or groupings (e.g., offender cohorts, offender type, districts, etc.) one variable (or outcome) at a time. The major characteristics or descriptive statistics to be examined for each variable include:

- Distribution (frequency distribution, percentages)
- Central tendency (mean, median, mode)
- Dispersion (e.g., range, variance, standard deviation).

These descriptive statistics will provide system staff and stakeholders with basic information about the offender population as a whole. [9]

Testing the "thread" of the logic model shown above, the following sample questions can be answered using univariate analysis:

- What are the demographics (e.g., distribution by race, age, gender, risk level, seriousness of previous criminal offense(s)) of the offender population entering supervision with an identified substance abuse problem?

- What percentage of offenders with an identified substance abuse problem participate in substance abuse treatment? What is the average dosage of substance abuse treatment received across this offender population? What is the most common modality of treatment provided?

- What percentage of this offender population abstains from using illegal or other restricted substances during the period of supervision?

- For those offenders with a history of substance abuse who are identified as using substances during supervision, what are the most common drugs of choice identified?

- What percentage of this offender population commits a new offense during supervision? Following release from supervision?

- What is the average length of time to a new offense for this population?

Additionally, by plotting the data using scatter-plots, histograms, bar graphs, or other options, it will be possible to identify potential differences or variation in outcome responses for subgroups of the offender population for further analysis. For example, are there visible differences in the distribution of offenders with a history of substance abuse with a new arrest while under supervision? Are there differences in the distribution of this offender population by reasons for ending supervision (e.g., successfully completed on time, early release, revocation, absconded)? Is there wide variation in the amount of treatment hours provided for this population of offenders?

Bivariate Analysis

The next step in the analysis is to begin testing the significance of the covariates (inputs, process (activities), process outcomes) on the intermediate (e.g., substance use, employment, life skills) and ultimate outcomes (criminal activity, victim restoration, compliance with release conditions) using bivariate analysis techniques, such as correlations, cross tabulations with chi-square statistics, t-tests, and analysis of variance. These techniques are recommended because they are relatively easy to conduct and provide straightforward interpretations. [10] The bivariate analysis will provide preliminary indication of the relationship between independent and dependent variables to be further tested with the inclusion of control variables (e.g., offender characteristics, offense characteristics, sentence characteristics, etc.) with more advanced analysis.

For example, analysis of variance (ANOVA) is used to test differences between two or more means. ANOVA helps to determine which variables have a significant influence on an outcome, and/or how much of the variability in the outcome or dependent variable is attributable to each factor. The two types of ANOVAs that can be used to test such differences include:

l One-Way ANOVA. A one-way analysis of variance is used when the data are divided into groups according to only one factor (e.g., level of risk, type of correctional programming, compliance status). The questions of interest are usually: Is there a significant difference between the groups? and if so, which groups are significantly different from which others? Statistical tests are provided to compare group means, group medians, and group standard deviations. [11]

- Multifactor ANOVA (MANOVA). When more than one independent or control variable is present and the factors are crossed (e.g., instant offense by prior record), a multifactor ANOVA is appropriate. Both main effects and interactions between the factors may be estimated.

Using the substance abuse example, a one-way ANOVA might be conducted to test differences in mean number of substance abuse relapses between those offenders with a history of substance abuse who participate in individual substance abuse treatment and those who participate in other modalities of treatment (group, family, and other). [12] A detailed example and interpretation of the results from an ANOVA are presented in Attachment 7.

Another useful bivariate analysis to be considered is the negative binomial regression model. This technique can be used to test whether different subgroups of offenders commit more frequent acts of noncompliance with conditions of supervision, relapse, and/or new offenses or other outcomes than other subgroups. Negative binomial regression models were developed specifically for the kind of distribution of failures that are likely to be observed with these offender data (i.e., a large portion of the offenders will not fail at all during the time observed, some will fail once, fewer will fail twice, and a handful will fail more often). This type of skewed distribution (if present) would violate the normality assumptions of ANOVA. [13]

Advanced Analysis

Based on the results of the univariate and bivariate analysis, more advanced statistical techniques (logistic regression, survival or Cox regression) are recommended to test the relationship between process outcomes, intermediate outcomes, and ultimate outcomes after controlling for inputs (e.g., characteristics, sentence, and resources) and process (system activities). The recommended advanced statistical techniques are described below.

Logistic Regression

Logistic regression allows one to predict a discrete outcome, such as group membership, from a set of variables that may be measured at any level (interval, ratio, ordinal, or nominal) or a mix of levels. Generally, the dependent or outcome variable is dichotomous, such as presence/absence or compliance/noncompliance. Logistic regression calculates the probability of success over the probability of failure, presenting the results in the form of an odds ratio. For example, logistic regression can tell us the probability that an offender will reoffend after controlling for various input or process factors. Logistic regression also provides knowledge of the relationships and strengths among the variables.

For example, it is possible to use logistic regression to test the relationship depicted in the logic model between reduced substance abuse and minimized criminal activity beyond the period of supervision. The independent variable would represent a measure of substance abuse during supervision. This could include a dichotomous variable representing presence or absence of relapse created from existing measures of substance abuse or a continuous variable representing the number of relapses during supervision. The dependent variable would represent criminal activity following supervision. For logistic regression, this would be represented by a dichotomous variable, such as presence or absence of a new offense, substance abuse-related or other type of crime, or other dichotomous measures of ultimate outcomes presented in the logic model. Additionally, it is important to include measures of inputs as control variables into the regression equation, such as characteristics of the offender, offense, and sentence. For ease of interpretation and to best understand the variance accounted for by the control variables, it is recommended that the inputs or controls be entered first in the equation as a block or set of control variables.

Unlike linear regression, the interpretation of the coefficient for logistic regression (Exp(B)) is more straightforward, representing the likelihood of an event occurring. Using the substance abuse example, the results of logistic regression can tell you how much more likely an offender who relapses during treatment is of committing a new (non-substance abuse-related) offense during and beyond supervision than an offender who does not relapse during treatment. With logistic regression, it is also possible to control for a block or set of variables (inputs, process, etc.) in the analysis.

Survival Modeling

Survival modeling is recommended when you want to examine the relationship among independent variables or covariates and the time to events of interest, for example, time to employment, time to relapse, time to recidivism, time to completion of paying restitution, etc. Static models alone are insufficient in this situation because they assume that a rearrest or other outcome is the same regardless of whether it occurred on the first or last day of the period of interest. The timing of these events, however, is a particularly important distinction when considering the public policy and safety implications of supervision. Survival modeling or analysis is an effective statistical technique to use when you want to examine the impact of time-varying covariates on these events. It is a particularly useful technique when comparing groups with varying follow-up periods. Survival analysis handles time at risk by subdividing the follow-up period into smaller observation points. At each of these points, the proportion of the sample that is at-risk for reoffending, for example, is used to estimate the probability of surviving beyond that point. This method ensures that only the characteristics of the population still at risk are used to estimate the time until failure, thereby providing a more accurate prediction of failure. Additionally, survival functions use maximum likelihood techniques that can differentiate between censored and uncensored cases. That is, survival modeling accounts for those cases who survive throughout the follow-up period. [14]

An example of the results of a hypothetical comparison (based on “dummy” data) of the risk of recidivism among offenders with a history of substance abuse who participated in individual treatment compared to offenders who participated in group treatment, after controlling for race, age, and risk assessment is shown in Attachment 7. Like logistic regression, the results of survival modeling are interpreted as the likelihood of an event occurring.

Linear and Multiple Regression

While logistic regression and survival modeling are valuable statistical techniques to use when the outcome or dependent variable of interest is a dichotomous variable, other regressions should be used when the outcome of interest is a continuous variable. For example, if you want to determine whether there is a relationship between offenders who demonstrate a reduction in substance abuse during supervision and the amount of restitution an offender is able to pay, linear regression should be used. If you want to examine the relationship of more than one variable, for example substance abuse, employment, and life skills, on an outcome (amount of restitution paid or number of new offenses), you should use multiple regression. These are all statistical techniques that will enable system staff and stakeholders to further explore threads of the post-conviction supervision logic model and identify predictors of success.

Trend Analysis

Trend analysis can be used to examine changes in outcomes over time for a given population, as well as to compare trends in outcomes across subgroups. Trend analysis is often depicted by a graph. This graph depicts hypothetical trends in the behavior of offenders entering post-conviction supervision with an identified substance abuse problem, specifically compliance with restrictive and financial conditions (represents percentage in compliance) over a 12-month period. The graph shows that a greater percentage of offenders were in compliance with restrictive compared to financial conditions during the 12-month period. However, both represent a gradual increase in the percentage of offenders in compliance from January to December. To predict future values for a variable, for example compliance with restrictive conditions beyond December, time series analysis is required. [15]

Analysis Assumptions

It is important to recognize that there are two key assumptions underlying the above analysis plan. The first assumption is that all independent and control variables are treated as exogenous variables. That is, a variable whose variability is assumed to be determined by causes outside the model under consideration. No attempt is made to explain the variability of an exogenous variable or its relations with other exogenous variables. Stated another way, none of the independent or control variables are said to affect (or to be affected by) any of the other variables. It is, however, recognized that these variables may be correlated with one another. The second assumption is that the logic model (multistage model) or segments of the model (single-stage model) being tested are well specified. Under this assumption, for regression analysis, it is possible to interpret the regression coefficient as the expected change in the dependent variable associated with a unit change in the variable in question, while partialing out the influence of the other variables (independent and/or control). [16]

As the logic model and hypothesized relationships are further developed and tested, and data collection refined, additional analysis (e.g., differences of proportions, interrupted time series) will be considered for future assessments.

5. Reporting

The results-based management framework will generate an informative annual research report that effectively shows and explains changes in outcomes over time and reasons or predictors of those changes. While the specific layout of the information in the report will be driven by the type of analysis and research questions to be addressed, a general template for the report is presented in Attachment 8 [omitted here]. Each section of the report is described in detail below.

I. Overview. This section will contain a standard description of the results-based management framework, including the overall purpose and a description of the logic model for post-conviction supervision underlying the framework.

II. Description of methodology. This section will contain a detailed description of the methodology behind the framework. In particular, it will include the research questions to be addressed, data sources and measures examined, selection criteria and resulting sample, and limitations to sample (e.g., limitations to generalizability of results to entire offender population).

III. Data quality and reduction process and results. This section will include a description of the data quality process (see Subsection 4.1) and the results of the analysis, including percentages for missing data, data errors, and other exclusionary factors that result in a reduction of the data set; description of data reduction methods (see Subsection 4.2) and results; and a discussion of any implications or limitations to the analysis as a result of the data quality assessment and data reduction process.

IV. Model fit. Because of the importance of the logic model to the integrity of the results, it is important to present the model fit results (test of the model as a whole) separate from the results for specific relationships depicted in the model. This section will include a description of the analysis conducted (e.g., principal component, path analysis, structural equation modeling), results (goodness of fit statistics), and an interpretation of the results. The interpretation is important for guiding the subsequent analysis and results. For the initial annual assessment, a complete test of the model may not be possible due to unavailable data.

V. Post-conviction supervision results. This section is critical to understanding what aspects of post-conviction supervision are producing desired outcomes and which areas need modifications or improvements. This section will be divided into two parts: ***demographics*** and ***outcomes*** . The ***demographics*** subsection should begin with the results of the univariate or descriptive analysis in order to provide the reader with a profile or profiles of the offender population, the system itself (districts, offices, regions, etc.), and other contextual factors that are important for interpreting the results. Additionally, it is important to point out that these measures will be used as control variables in the advanced analysis of outcome measures. The results of the bivariate analysis should be presented next to identify and provide evidence for testing specific relationships among processes and outcomes.

The next subsection will present the ***outcomes*** for post-conviction supervision. The logistic regression should be presented first with a description of the analytic technique (see Subsection 4.3), followed by an interpretation of the results for each run. This should include a description of the control variables, the relationship(s) being tested, and the findings. Where appropriate, graphics should be used to present the findings. Next, the results of the linear and multiple regressions should be presented. Together, the findings from the logistic and linear/multiple regressions will provide important information regarding the specific relationships depicted in the logic model (e.g., does X lead to Y when controlling for A and B?).

The regression results will be followed by a presentation of the survival modeling. A description of the analytic technique (see Subsection 4.3) will be necessary to ensure the reader understands why survival modeling is being used. In particular, it is important to explain that survival modeling is needed to compare outcomes for the different entering offender cohorts compared annually. Additionally, the results of the survival modeling can help predict likelihood of success (or failure) for an offender in the absence of complete follow-on data for all offenders. This will be important for the first several years of the assessments. Again, the use of graphs and charts to present findings is recommended. The results and an interpretation of each run will be provided.

Finally, the results of the trend analysis will be presented. This information will provide the reader with a description of how outcomes have changed over time and for which offender populations. Additionally, the effectiveness of different treatments or interventions can be compared over time. Using line graphs is the most effective method for presenting trend data. An interpretation of the findings will be presented.

VI. Implications and Recommendations for Policy and Practice. This is the most important section of the report. It will begin with a summary of key findings and a discussion of any unexpected findings and limitations of the data. Next, implications of the findings for setting priorities and making policy, programming, and resourcing decisions need to be presented followed by specific recommendations *supported by the results* . If possible, suggestions for implementing the recommendations should also be included.

Subsequent reports will follow the same template but will need to address, if appropriate, the following:

- Changes to data collection (process, sources, measures) and explanation for changes
- Changes to logic model and explanation for changes
- Changes to research questions (or focus of analysis) and reasons for changes.

It is important to include the core sections in each annual report but recognize the content may change and the format may need to be flexible. For example, once pretrial information is incorporated into the framework, it may be necessary to create a separate report template to present these findings or make modifications to the existing template in order to combine the results into a single report. Feedback should be obtained from end-users and changes made to the report template as appropriate to ensure clear communication of results and the usability of information.

* * *

7. Institutionalizing the Framework

Once the framework has been implemented, it is important to ensure it maintains momentum and continues to receive attention by management. Success can be measured by the extent to which the results-based management framework becomes institutionalized within OPPS. Just as the 2004 Strategic Assessment of the Federal Probation and Pretrial Services System recommended the need to organize, staff, and resource to promote mission-critical outcomes, the same can be said for the sustainability of the framework needed to assess progress toward those outcomes. Accomplishing this task will require adherence to a predefined yet flexible process, strong management support, and a commitment of resources. Each of these critical factors is described in the subsections below.

7.1 Predefined Process

To ensure full implementation of the results-based management framework, a continuous process of assessment, review, and modification is necessary. The specific steps in the process are outlined below.

Ongoing Consistent Assessments

Ongoing assessment is defined as annual extraction, analysis, and reporting on processes and outcomes for offender cohorts entering post-conviction supervision during specific fiscal years. Over time, the plan is to analyze data for three consecutive entering cohorts at a time. It is important to adhere to the criteria for selection each year in order to provide comparisons over time and to identify trends in outcomes. It is critical to ensure comparison of “apples to apples” each year and over time. Additionally, the timing of the extraction and reporting must be consistent each year…. To ensure ongoing assessment and the production of results that can be used by managers to make important decisions will require a review of existing personnel to identify those individuals with the necessary skill-set to conduct the assessments (e.g., knowledge of PACTS, ability to apply advanced statistical techniques and interpret results as identified in the analysis plan, etc.).

As baseline data become available for the various components of the logic model, it will be possible to set performance benchmarks against the baseline measures. While baseline measures are indicators of where the system is, benchmarks identify where the system needs to be in the future. It is important to use data (baseline measures) and actual experience when setting benchmarks to ensure they are realistic. This process should be a collaborative effort involving input from the field. Once benchmarks are set, they should be reexamined at least every three years to assess progress and determine if modifications are needed based on the results.

Review and Modify (Feedback loop)

The success of the results-based management framework relies on quality data, appropriate analysis and interpretation of results, and the utilization of the results. In particular, using the results to review and modify the framework is important to the longevity of the model.

As policies and practices change, it may be necessary for the model to change. As with the benchmarks, it is important to review the framework design, process, and results at least every three years to identify any needed changes or modifications to the logic model, measures, data systems, selection criteria, etc. Additionally, information needs of management and the field may shift, requiring changes to the framework. This feedback loop will help ensure a results-based management framework that is responsive to changes over time. Any modifications to the system need to be vetted through key stakeholders, including the field.

7.2 Management Support

It is critical for the results-based management framework to be owned by a specific organizational unit within OPPS. OPPS may wish to consider restructuring the existing organizational structure to create a new unit focused exclusively on results management and the implementation and sustainability of the framework. Whatever approach is taken, it is important that there be a manager whose primary responsibility is the oversight of the framework. The manager must also have the authority or access to the appropriate lines of authority to ensure support of the framework and consideration of recommendations to decisions regarding policy, programming, and resourcing. To the extent possible, it is also preferable that the unit is viewed as independent of the other divisions and branches. This is important to ensure objectivity, neutrality, and ensure the unbiased reporting of findings and recommendations. Additionally, sustainability of the framework requires designated staff with the expertise necessary to ensure data quality, analyze data, and translate results into practical information. Staff also need to be able to make modifications to the framework, including revising the logic model and identifying and operationalizing new measures.

7.3 Commitment of Resources

As with any new effort, it requires resources to get a process up and running and to continue operating over time. A thorough assessment of the resource needs for completing the remaining implementation tasks and sustaining the framework by carrying out the lifecycle plan needs to be conducted. This assessment should be reviewed annually, especially within the first three years as changes and modifications requiring additional resources are likely. It is clear that staff and resources need to be organized to support the framework.

8. Next Steps

This lifecycle document is intended to serve as the primary document that describes the content of the results-based management framework, the analytic approach to the data, implementation process, and plans for institutionalizing the framework within OPPS. It is important that the information presented in this document be reviewed by system staff and stakeholders in order to verify the information, fill in gaps, review recommendations, and resolve unanswered questions. Once the framework is fully implemented, it will provide system staff and stakeholders with the information needed to better assess performance—and make programming and resourcing decisions—based on what the federal probation and pretrial services system accomplishes rather than solely on what it does.

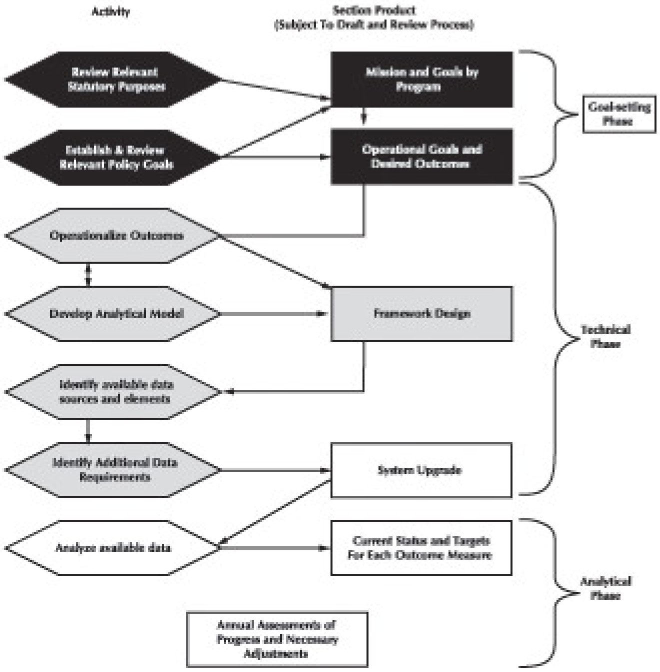
back to top

Endnotes

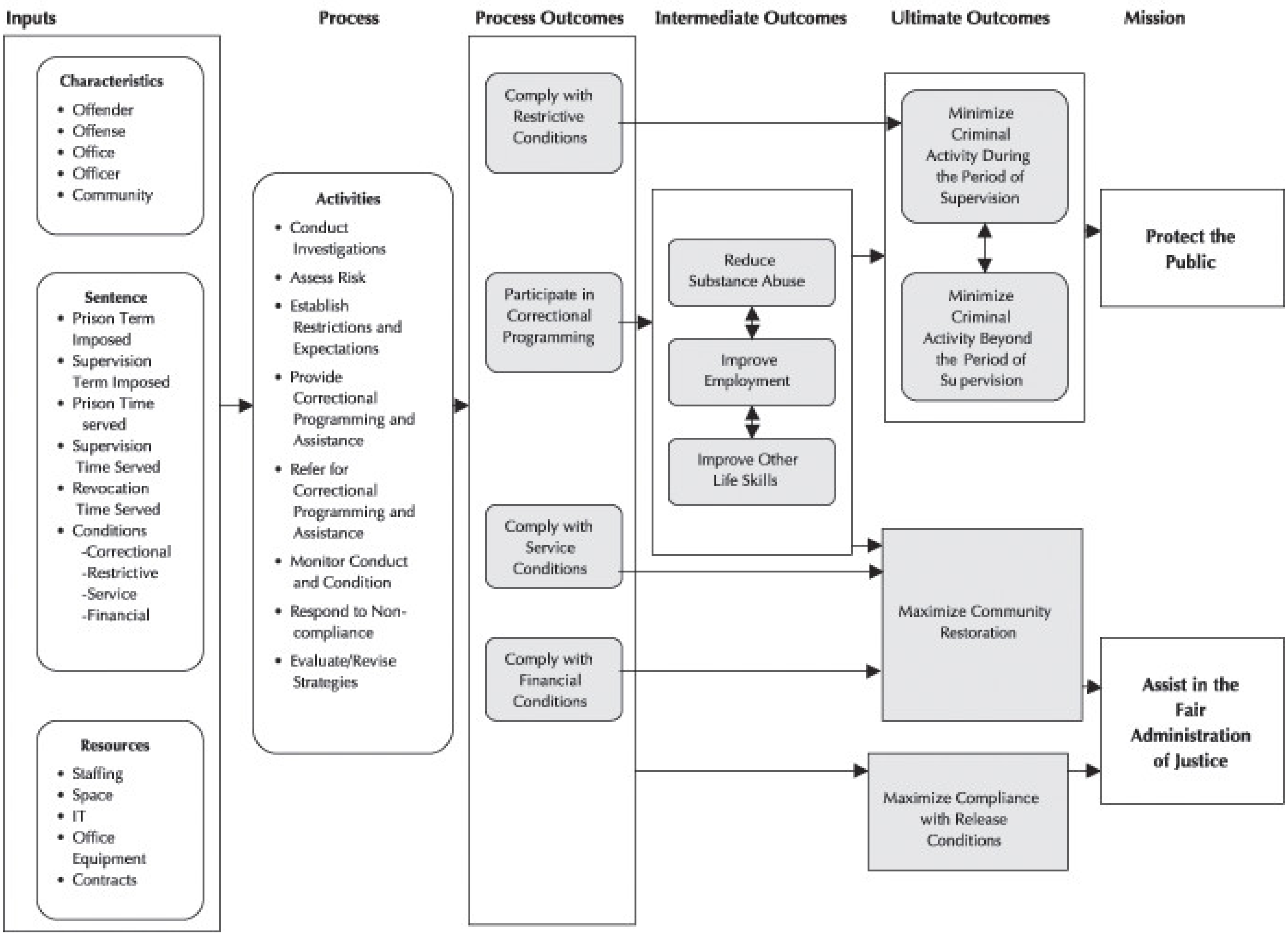
Published by the Administrative Office of the United States Courts www.uscourts.gov
Publishing Information

STATUS OF PROCESS FOR ESTABLISHING OUTCOME INDICATORS FOR SUPERVISION FUNCTIONS

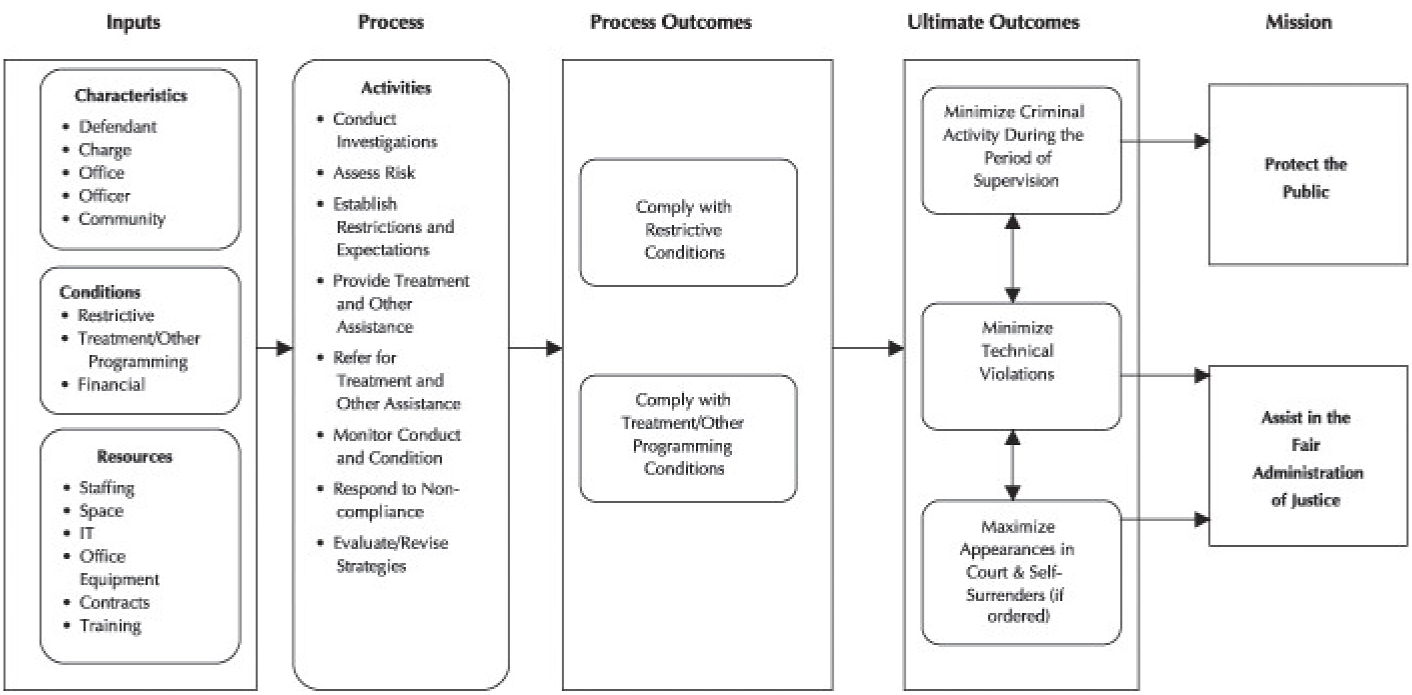
(Black=Completed; Gray=In Progress; White=Future Task)

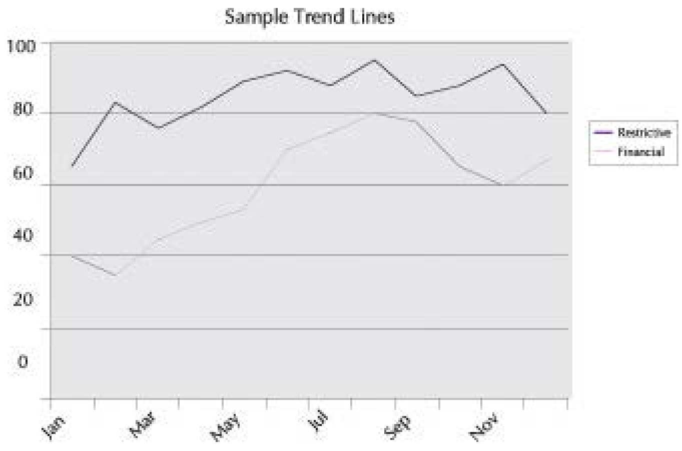


Post-Conviction Supervision



Attachment 4: Pretrial Services Supervision Logic Model




Sample Trend Lines
100
80
60
40
20
0
Jan
Mar
May
Jul
Sep
Nov
Restrictive
Financial

representatives in the sample for some combinations of gender, race/ethnicity, court type, and age.

[6] This includes those with negative scores.

[7] We include the square of age as a factor in the logistic regression because age has a curvilinear relationship with rearrest. Little (n.d.) used a similar analytic approach in her evaluation of the SDRRC.

back to top

How Much Risk Can We Take? The Misuse of Risk Assessment in Corrections

[1] See http://bmj.bmjjournals.com/cgi/content/full/312/7023/71. Also see http://www.ahrq.gov/clinic/epc/ for a listing of the growing number of evidence-based practice medical and mental health centers in the U.S. and Canada.

[2] Lowenkamp, Christopher T. and Edward J. Latessa. 2005. "Evaluation of Ohio's CCA Funded Programs. Final Report." Cincinnati, OH: University of Cincinnati, Center for Criminal Justice.

[3] Austin, James. 2006. "What Should We Expect From Parole?" American Probation and Parole.

[4] For a recent summary of the validity of the LSIR and its history see Girard, Lina and J. Stephen Wormith, 2004. "The Predictive Validity of the Level of Service Inventory- Ontario Revision on General and Violent Recidivism among Various Offender Groups" *Criminal Justice and Behavior*, Vol. 31, No.2:150-181 and Violent Recidivism among Various Offender Groups" *Criminal Justice and Behavior*, Vol. 31, No.2:150-181.

[5] For more information on MHS, Inc see their website at http://www.mhs.com/index.htm.

[6] For more information about Northpointe see their website at http://www.northpointeinc.com/contact.htm.

[7] Washington State Institute for Public Policy. *Washington's Offender Accountability Act: An Analysis of the Department of Corrections' Risk Assessment.* December 2003. Olympia, Washington; James Austin, Dana Coleman, Kelly Dedel-Johnson, and Johnette Payton. 2003. *Reliability and Validity of the LSI-R Risk Assessment Instrument* . Washington, DC: The Institute on Crime, Justice and Corrections at the George Washington University; and James Austin, 2006. *Vermont Parole Board Risk Based Guidelines, Technical Assistance Report #2*. Washington, DC: National Institute of Corrections.

[8] Washington State Public Policy Institute. 2003. p.4

[9] Austin, James, Dana Coleman, and Kelly Johnson. 2002. "Reliability and Validity of the LSI-R for the Pennsylvania Board of Parole and Probation." Washington, DC: The Institute on Crime, Justice and Corrections, The George Washington University.

back to top

From Theory to Practice: The Lifecycle Document for the Results-Based Management Framework for the Federal Probation and Pretrial Services System

[1] The terms "outcome-based" and "results-based" are used interchangeably.

[2] The term "evidence-based practice" implies that 1) there is a definable outcome(s); 2) it is measurable; and 3) it is defined according to practical realities (e.g. recidivism).

[3] As shown by the framework development flow chart, the process is iterative. All references to "completion" refer to the initial development process.

[4] While the focus of the measurement matrix is on outcomes, examples of measures for the inputs depicted in the logic model have been provided. It is expected that these will be revised and refined based on review and comment from system staff and stakeholders.

[5] While most of the imputation-based procedures described are available through SPSS, other more advanced procedures may require additional statistical software, such as SAS or Stata.

[6] For principal component analysis, variance maximizing (varimax) rotation should be selected as the extraction method. Additionally, different criterion (Kaiser criterion, scree test) should be examined to determine which solution makes the best sense, often one retaining more factors (Kaiser) than the other (scree).

[7] There are over 25 goodness-of-fit calculations available through the SPSS add-on AMOS. The most common used are model chi-square (not significant indicates model fit), GFI (goodness-of-fit index) (greater than or equal to .90 to accept the model), and CFI (comparison-fit-index) (greater than or equal to .90 to accept the model). Any or all of these should be compared.

[8] Path analysis and structural equation modeling can be conducted using SPSS and the SPSS add-on software AMOS as well as SAS. necessary to modify the model (add and/or remove arrows depicting relationships, add/or remove variables, etc.).

[9] Most univariate analysis can be conducted using the Descriptive Statistics option within the Analyze function of SPSS.

[10] Bivariate analysis can be conducted by using the Analyze function of SPSS and selecting Descriptive Statistics, Compare Means, and Correlate options.

[11] When comparing means using ANOVA, multiple range tests are used, the most popular of which is Tukey's HSD procedure.

[12] ANOVA can be conducted by using the Analyze function in SPSS and selecting the Compare Means/One-Way ANOVA option.

[13] This particular statistical technique is not available through SPSS. It would require the use of Stata statistical software.

[14] Survival analysis can be performed using the Analyze function and selecting the Survival/Cox Regression option in SPSS.

[15] While you can use SPSS to generate trend graphs, a software add-on, SPSS Trends, is required to conduct more sophisticated time series analysis.

[16] When comparing across different populations, it is important to use b's (regression coefficients) rather than $'s (standardized regression coefficients) because they are more sensitive to fluctuations in variances and covariances across populations.

back to top

References
