
Correctional Operations Trend Analysis System (COTAS): An Independent Validation

Submitted to:
The Florida Department of Corrections

Prepared by:

Florida State University
College of Criminology and Criminal Justice
Center for Criminology and Public Policy Research

Thomas G. Blomberg
Jim Clark
Leslie Hill
Bill Bales
Karen Mann

Completed June 30, 2011

Chapter 1

Overview of the COTAS Validation

Introduction

This report provides research findings and recommendations from the Florida State University (FSU) Center for Criminology and Public Policy's validation of the Florida Department of Corrections (DOC) Correctional Operations Trend Analysis System (COTAS). COTAS is designed to serve as a tool for DOC staff to aid in the prevention of violent events at institutional facilities. Using DOC's large collection of historical and real-time data regarding characteristics about individual inmates, violent and non-violent incidents, and environmental characteristics of institutions, COTAS provides correctional administrators with trend analysis and risk assessment of inmates' involvement in violent events.

COTAS provides the regional and facility administrators with two types of statistics: descriptive and predictive. Descriptive statistics from COTAS provide a summary of violent and non-violent events that occurred within DOC regions or administrative areas within the prior 30 days. This data can be "drilled down" to examine the prevalence of events at the facility, dorm, and inmate levels. Additionally, COTAS can provide a 12-month trend analysis of facilities' monthly count of specific violent and non-violent events. Predictive statistics from COTAS provide the predicted probability of individual inmates' involvement in violent events. Predictions are generated by an algorithm, which uses historical data to examine the relationship between inmate and facility characteristics and inmates' involvement in violent events. Both descriptive and predictive statistics are reported to the user in a web-based dashboard interface. Based on pre-defined thresholds, the interface (dashboard) displays the degree of concern that a particular administrator should have regarding the likelihood of violent events occurring during the next thirty days. A detailed description of the COTAS system is provided in Chapter 2 of this report.

The following sections of this chapter provide a brief history of the conception and development of COTAS followed by a description of the validation methodology employed by project staff. This Chapter concludes with an overview of the remaining chapters in this report.

History of COTAS

In 1999, COTAS became a concept in DOC's Bureau of Research and Data Analysis. It evolved into a two-level model that determines: (1) inmates that are most likely to be disruptive; and (2) facility characteristics that are predictive of disruption. At that time, DOC was unable to obtain the resources or technology to fully develop and implement

the project, however it continued to evolve as a concept until grant funding became available.

In 2006, DOC was awarded a grant of \$500,000 from the National Institute of Justice (NIJ) to develop COTAS. The NIJ award proposed to develop a software system that creates a correctional “crime mapping and information management system to monitor cross functional operations that can be quickly viewed by all levels of correctional management.” It was designed to identify trends and patterns in violent incidents within institutions and statewide. Thus, correctional staff and administrators would be empowered with information allowing them to take actions to attenuate or curtail violent events.

DOC identified four separate rationales for reducing the occurrence of violent events in facilities. First, limiting the occurrence of violent events has the direct effect of enhancing the safety of staff and inmates. Second, reducing the occurrence of violent events is cost effective. For example, it has been estimated to cost \$970 per infraction (both violent and non-violent) at a medium security institution (Lovell & Jemelka, 1996). Third, the reduction of violent events will minimize the interruption of the prison adjustment period for inmates. During periods of unrest, inmates are not able to take full advantage of rehabilitation programs. Additionally, DOC stated that inmates with paying jobs may have their pay cut due to the prison being put on lock-down during extreme disorder. Pay reductions will interrupt their restitution effort. Finally, the reduction of violent events in prison, such as rapes, assaults and homicides, will limit the long-term effects on correctional staff, inmates and the community.

The first Project Director for the COTAS development project was S. Fred Roesel, former Chief of the Bureau of Classification and Central Records. David Ensley, Chief of the Bureau of Research and Data Analysis (BRDA), and John Agliato, former Chief of the Bureau of Systems Development, served as the Research Director and the Information Technology Director, respectively. At the time of this validation, David Ensley is the only member of the leadership team involved with COTAS. Development of COTAS was primarily split between two offices: BRDA, which identified the predictors of violent events and the Office of Information Technology (OIT) which created the extract, transfer, and load (ETL) processes to transfer real-time data from DOC’s mainframe databases to a central data warehouse and creates the user interface of COTAS.

BRDA limited the selection of predictors of violent events to those that were “well-defined, easily understood by users, and significantly predictive of disruptive events.” Key data elements were identified using five years of monthly snapshot data of inmate, facility, and staff characteristics as well as input from prior literature and correctional administrators. The statistical package SAS was used to determine the best inmate predictors of violent events by facility category through regression analysis. Checks of multi-collinearity were conducted to remove duplicative or highly-correlated predictors. A final list of significant predictors was provided to the programmers for the development of the predictive models.

In October 2007, DOC contracted with Idea, a business intelligence consulting and design firm, to develop a data warehouse, COTAS applications, and COTAS interface. A single data warehouse was designed with a multidimensional schema to store data. The data warehouse is housed on a server running Microsoft SQL Server 2005 database software. Data was pulled into the data warehouse from a number of DOC databases including:

- OBIS (Offender Based Information System): a centralized mainframe hierarchical data store used to maintain and record offender/inmate records;
- HRD (Human Resources Database): a centralized server based data store used to maintain and record Department staff records;
- FAST (Facility Access Secure Tracking): a centralized and distributed server-based data store to maintain and record the occurrence of inmate visitation and inmate volunteer activity;
- IGLOGS: a centralized server-based data store used to maintain and record investigative data by the Department's Inspector Generals Office and inmate grievance data;
- MINS (Management Information Notes System): a centralized mainframe hierarchical data store used to maintain and record staff/offender incidents throughout the Department;
- Inmate Gang Database: a centralized server-based data store used to maintain and record inmate/offender gang activity;
- Use of Force Database: a centralized server-based data store used to maintain and record whenever a staff member uses force against an inmate or offender.

The data warehouse serves as the central data source for the calculation of both COTAS descriptive and predictive statistics and is updated with real-time data Monday through Thursday evenings. The Microsoft SQL Server 2005 database software platform was used because it allowed for the integration of Microsoft Data Mining software and could be purchased and maintained at a lower cost to DOC relative to other data mining software. A detailed description and analysis of the data warehouse, ETL processes, and user interface is provided in Chapter 6 of this report.

In April 2008, an initial rollout of COTAS to a test group was conducted to garner feedback and recommendations for improvement. The results of this initial test were used to enhance the functionality of COTAS. In June, 2008, the first production version of COTAS was released to prison administrators. In October 2008, DOC obtained additional grant funds (\$122,070) from NIJ and COTAS was enhanced based on input from the test group. Modifications to COTAS were also implemented as a result of a meeting of the Warden's Workgroup in April 2009. The Workgroup was composed of five wardens from different facility types (e.g., reception centers, work release programs, secure facilities). In 2009, DOC received additional funding, a \$100,000 award, from

NIJ to continue the development of COTAS, to improve transferability of COTAS to other state correctional initiatives, develop documentation/manuals, and to fund an independent validation. DOC received an additional \$150,000 from NIJ to complete these tasks. In 2010, DOC issued a “Request for Quote” seeking applicants to conduct the validation. In February, 2011, DOC entered into a contract with the Center for Criminology and Public Policy at FSU to conduct the independent validation of COTAS. The next section provides an overview of the COTAS validation methodology.

FSU conducted the validation of COTAS between February 1, 2011 and May 30, 2011, with the final report released June 30, 2011. In most instances, an independent evaluation or validation would occur in a separate environment—one that is independent of the sponsor or funder of the evaluation or validation. An independent or neutral environment is most desirable and one which increases the likelihood of objectivity and minimizes the risk of the evaluator being co-opted or influenced by the sponsor or funder. In the spirit of full disclosure, it should be noted that DOC required FSU’s project staff to work within DOC offices in order to conduct this validation. Therefore, FSU staff were housed in and worked directly from DOC offices during the validation process of COTAS. FSU contracted with a highly qualified firm to conduct the software review and validation; these professionals visited DOC to meet staff, receive an overview of the system, and conduct an initial review of COTAS; however, they were provided remote access and were able to carry out the tasks of the software validation within an independent environment external to DOC.

Overview of the COTAS Validation

The COTAS validation focused on three components: (1) the predictive measures for violence and the thresholds used to distinguish levels, (2) the programming and software implementation, and (3) the ability of COTAS to provide accurate, understandable, meaningful, and constructive information to end-users. The validation of the predictive measures for violence and the establishment of thresholds focused on two elements of predictive accuracy: (1) the appropriateness of the modeling procedure and (2) the discrimination of the predictive model. The appropriateness of the modeling procedure refers to the correct selection and application of the statistical model. For example, are all of the mathematical assumptions of the data satisfied? Discrimination refers to the model’s ability to distinguish between low-risk and high-risk events and, therefore, the ability to set significant threshold variations. For example, are inmates with high risk scores involved in a higher proportion of violent events relative to inmates with low risk scores?

The validation of the software implementation and aggregation of data was conducted by ELENCO, Inc, a private consulting and software design firm. The software validation procedures were divided into five key evaluation tasks:

- (1) Software design evaluation (e.g., how well is the software designed and coded?)
- (2) Software Implementation (e.g., how well is the application implemented?)
- (3) User Documentation (e.g., are the materials accurate, sufficient, and up-to-date?)

- (4) Software Transferability (e.g., can the software be transferred to and used by another agency in a cost effective manner?)
- (5) Software Extensibility (e.g., can new software features and functions be added for a reasonable cost?)

The software validation methodology and results are presented in Chapter 6.

An assessment of COTAS users' experience was conducted through a web-based user feedback survey, which examined users' experiences with and knowledge of the system and provides suggested improvements to COTAS. Surveys were administered to correctional administrators at all DOC facilities in the state.

Structure of the report

Chapter 2 of this validation report presents an overview of COTAS including a detailed description of the functions and the user interface. Chapter 3 provides an overview of the validation including a description of the methodology, an analysis of the model's ability to predict inmates' involvement in violent events, and the results. Chapter 4 examines users' experiences with COTAS and presents the findings of the survey of COTAS users. Chapter 5 summarizes the COTAS validation and presents recommendations that may improve the accuracy and usability of COTAS. Chapter 6 presents ELENCO, Inc.'s software validation report which includes an examination of the ETL procedure for the data warehouse and the software design of COTAS.

Chapter 2

Overview of COTAS

Introduction

As discussed in the previous chapter, the primary goal of COTAS is to reduce the incidence of violent events in facilities by providing prison officials with accurate, timely, and actionable information. Specifically, COTAS is designed to perform three tasks: (1) identify key predictors of future prison disruptions, (2) use real time data to make predictions, and (3) communicate this information using a clear and intuitive software interface to empower prison administrators to quickly and easily identify key risk factors for violent events within their facilities. COTAS is a complex data system with many dynamic attributes. Thus, it is helpful to provide a general overview of COTAS prior to discussing the validation of the system. The prior chapter presented the primary rationale for COTAS and a brief history of the development of the system. This chapter provides a detailed description of the functions and displays of COTAS. The first section presents an overview of key components of COTAS including a logic map of the structure of the system. The next section provides a detailed description of the COTAS functions and user interface, including screen shots and a guide to interpreting statistics displayed to users. The chapter concludes with a discussion of the system design and presents key recommendations for improving the overall system.

The Structure and Key Components of COTAS

COTAS is a joint venture project developed by DOC's BRDA and OIT. This collaborative effort was key in the development of COTAS because of the need for expertise in statistical analysis, database integration, and software design. BRDA focused first on identifying the measures that significantly predict future inmate involvement in violent events. A violent event was defined as violent disciplinary reports, violent escapes, investigations of violence, or reported use of force. BRDA analyzed five years of monthly snapshot data to determine the statistically significant predictors of inmates' involvement in a violent event. The "simplest and smallest number" of predictors was chosen for inclusion in the predictive model or algorithm. This list was determined based on the predictive accuracy of the measure, the malleability of the measure (e.g., could facility administrators take actionable steps to mitigate the odds of an event occurring?), and the parsimony of the overall model (e.g., how much did a predictor contribute to whether an event occurred?). The validation team was unable to locate documentation that tracks the historical development and decision points of COTAS in detail; therefore, a list of the variables that BRDA provided to the OIT and the consultant, *Idea* is not included in this report. DOC indicated that some of the variables that BRDA identified as highly predictive are included in COTAS, the final list was generated by *Idea*. It appears that staffing variables and gang variables from BRDA's list of predictors were excluded from the final list of variables. The variables included in the inmate level logistic regression include:

- involvement in a violent event (DV),
- inmates' age,
- gender (sex),
- bed category,
- number of months since an inmate's last violent event,
- number of prior violent events,
- positive drug test,
- race (white/non-white),
- time served,
- violent offender flag, and
- prior placement in a closed management bed category.

Additionally, some measures were removed due to concerns of autocorrelation. Autocorrelation occurs when two or more predictive measures are so strongly correlated with one another that they virtually capture the same underlying concept. For example, in research on academic achievement, childrens' *age* and *grade in school* are strongly correlated. Indeed, age is one of the major determinants of a student's grade in school and progression. Autocorrelation becomes problematic in statistical analysis because the correlated predictors in essence cancel each other out by each capturing the others' explanatory power. Thus, it is common practice for statistical analysis to leave out predictive measures when they are highly correlated with other predictors in the model or algorithm.

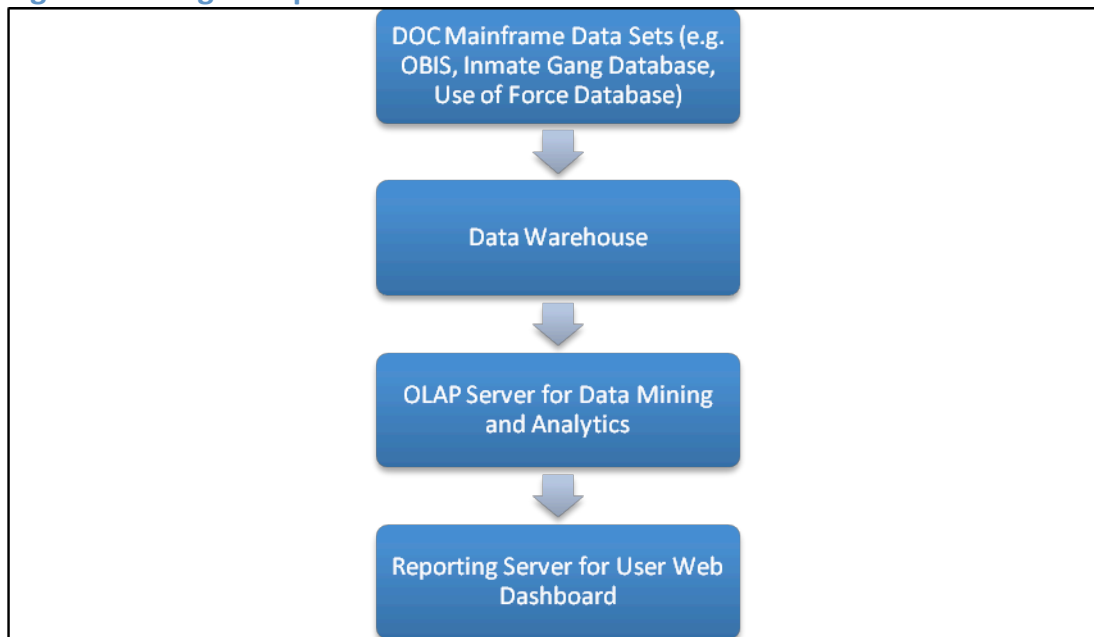
In addition to considering predictor accuracy, malleability, parsimony, and correlation, the predictive model for COTAS was developed to compare only like institutions by means of statistical controls because predictors in a youthful offender facility are likely different from predictors in a maximum custody facility. BRDA provided staff in OIT a list of measures to include in the algorithm for COTAS.

OIT had three main tasks in the development of COTAS. First, OIT developed the data warehouse (the repository of data used by COTAS) and the ETL procedures for populating and refreshing the data warehouse Monday through Thursday nights. Second, OIT created algorithms for COTAS' predictive models using some variables from the list of predictive measures from BRDA and variables deemed highly predictive by the SQL Server Analysis Services from Microsoft. Third, OIT developed the application that: (1) consolidates the measures into a web-based users "dashboard," (2) provides documentation for end users, and (3) provides transferability documentation to replicate COTAS in other states or governmental agencies. Additionally, OIT is responsible for the daily operations, maintenance, and development of COTAS. Many of OIT's programming tasks for COTAS were contracted out to *Idea*, a private consulting and development firm with a specialization in database warehousing and business intelligence software.

Figure 2.1 provides a general overview of the application logic of COTAS. Data are transferred from DOC's existing mainframe databases to the data warehouse. The data

from the data warehouse are used to run predictive algorithms and descriptive statistics. Results from statistical analysis are transferred to a reporting server and then used to create reports for users via a web-based user interface or dashboard. The next section of this chapter discusses the functions presented in the COTAS user interface.

Figure 2.1: Logic Map of COTAS



COTAS Functions and Users Interface

The COTAS web-based interface is accessible to users on DOC’s intranet allowing employees of DOC to have access to COTAS reports and links to other DOC websites. As previously mentioned, COTAS provides users with descriptive and predictive statistics. Additionally, the COTAS web-based interface allows users to explore aggregate-level statistics and individual cases. For example, a user may view the overall number of violent incidents in the past 30 days in an administrative area or facility, the different types of incidents in a facility, or the individual inmates involved in each type of incident. Additionally, the user interface provides links to other DOC resources and reports.

Figure 2.2 shows an example of the RegionGauges dashboard (the first screen upon logging into COTAS). The four gauges on the top-half of the screen display the number of facilities (by region) where violent events exceeded the established thresholds (a discussion of the thresholds will be discussed in the following section). The four pie charts or “lifesavers” at the bottom of the screen represent the proportion of inmates (by region) that are categorized as having a low (green), medium (yellow), or high (red) predicted probability of committing a violent event in the next month. The column on the left side of the screen allows users to: (1) access information on an individual inmate using the “DCNumber”, (2) access the Prison Rape Elimination Act (PREA) data, (3) view the

frequency of all violent and non-violent events by region, and (4) access high profile inmate reports by region and facility. Thus, the four gauges at the top half of the screen, and the links and data on the left-side column present descriptive statistics, whereas the four “lifesavers” at the bottom of the screen present predictive statistics. The section below provides an overview of the *descriptive* statistics in COTAS, followed by an overview of the *predictive* statistics available to users.

Figure 2.2: RegionGauges Screen



Descriptive Statistics

As stated above, the four regional gauges at the top of the RegionGauges dashboard (Figure 2.2) display the number of facilities in a specific region that are above a predetermined threshold for the number of inmates involved in violent events that occurred in the past 30 days. For example, in Region 1, two facilities were above the threshold for the number of inmates involved in violent events. The gauges on the top half of Figure 2.2 are color-coded with green (< 7 facilities in the region above the threshold), yellow (7-17 facilities in the region above the threshold), and red (18-25 facilities in the region above the threshold). This color-coding is arbitrary in terms of overall violence within a region. Indeed, the four designated regions contain varying numbers of facilities (between 34 and 41) thus the thresholds for the color designations have unclear substantive meaning because they may reflect the number of facilities in a district *and* the overall violence within a district.

Figure 2.3 provides an example of the RegionGraph screen. Users can navigate to the regional RegionGraph screen by clicking on the gauge regional title (e.g., Region 1) on the RegionGauges dashboard. The RegionGraph displays monthly variation in the number of violent events within a given region by the type of event (i.e., disciplinary report, investigation, use of force, and violent escape).

Figure 2.3: RegionGraph Screen

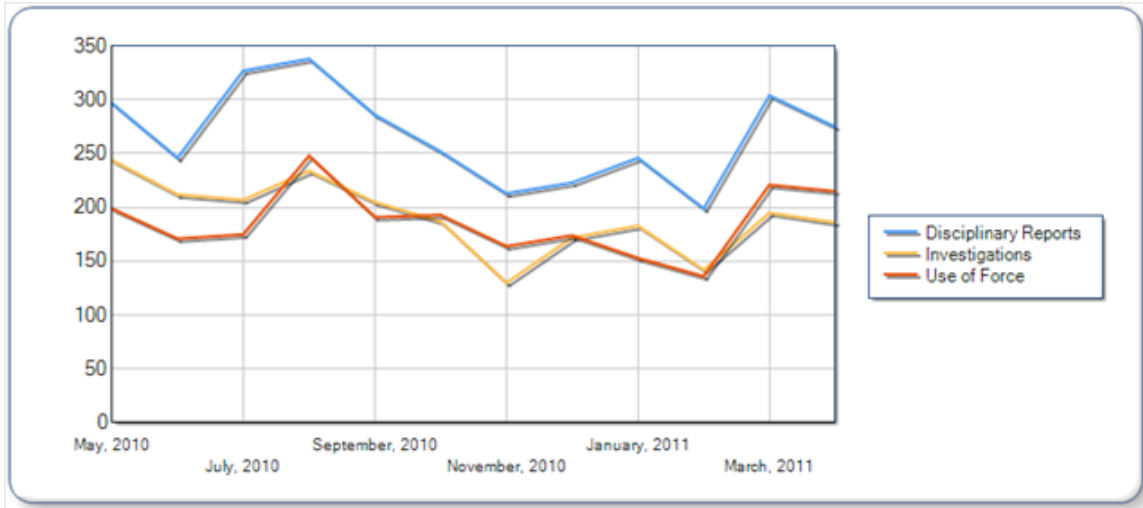


Figure 2.4 provides an example of the RegionDashboard screen. The RegionDashboard screen can be navigated to by either clicking on the region gauge at the top of the RegionGauges dashboard or by clicking on the region name on the left-side column of the RegionGauges dashboard. The RegionDashboard allows users to view data at the facility level. All of the facilities in the region are listed by color-coding, category, facility type, and number of events (both violent and non-violent). Additionally, a map of the region is displayed with the location of each facility and regional event information displayed on the right side of the screen.

Figure 2.4: RegionDashboard Screen

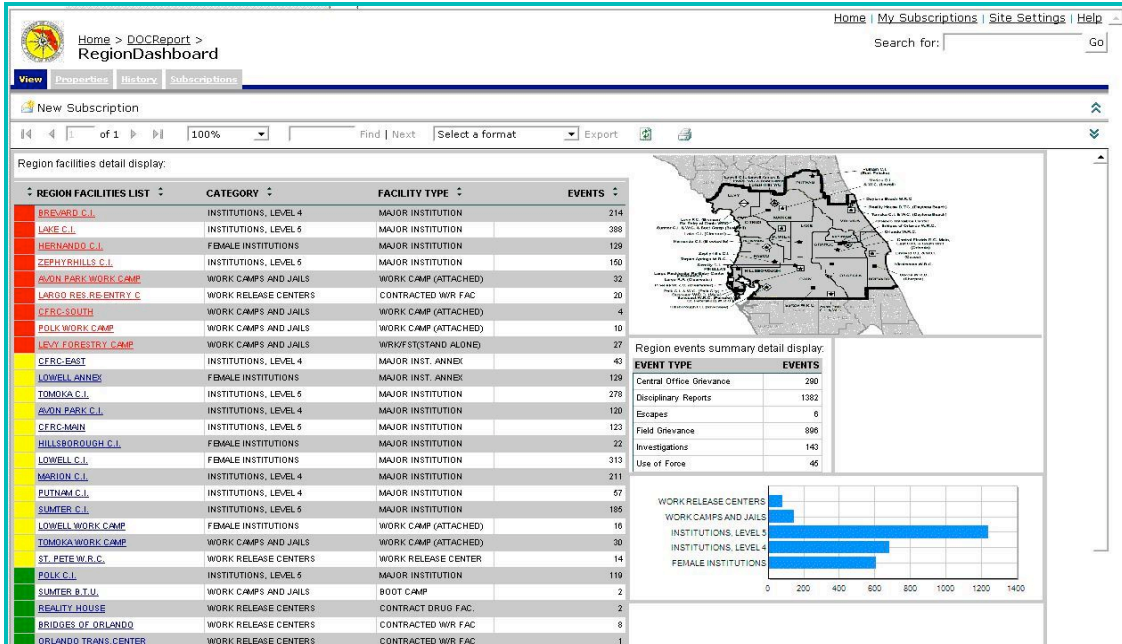


Figure 2.5 provides an example of the FacilityMonitor screen. The FacilityMonitor screen can be reached by clicking the name of an individual institution on the RegionDashboard’s list of facilities. The four gauges at the top of the screen present the number of violent events by the type of event (the number in the bottom center of the gauge) and the percentile rank that the number represents relative to historical data for similar facilities. This historical data was unavailable for analysis. The percentile rank determines the color-coding for the institution (green = ranking in <30th percentile, yellow = ranking in the 30th to 79th percentile, and red = ranking in the 80th percentile or above). There are two exceptions that will place institutions automatically in yellow or red categories. If an institution experiences a violent escape, it will automatically be placed in the “red” category¹. Also, if an institution falls into the “yellow” or “red” category for any of the four violent event categories (regardless of its percentile ranking), it will be coded as “yellow” or “red.”

The five gauges on the bottom of the FacilityMonitor screen present the number of non-violent events by the type of event. The number in the bottom center of the gauge is a count of the events and the dial on the gauge presents the percentile rank that the number represents relative to historical data for similar facilities. While non-violent events are color-coded in a similar manner to violent events; however, non-violent events have no impact on the color coding of the facility as a whole. Clicking on the labels and gauges of the FacilityMonitor screen allows users to explore annual trends in the number of

¹ COTAS codes all escapes as violent escapes; currently the system does not distinguish between nonviolent escapes (e.g., walk away from work release, perhaps the majority of escapes) and violent escapes.

monthly events and view specific information about the event including reports and inmates involved.

Figure 2.5: FacilityMonitor Screen



Predictive Statistics

In addition to descriptive data, COTAS also allows users to examine the predicted probabilities of individual inmates’ involvement in a violent event within the next 30 days (predictive statistics). The “lifesavers” on the bottom of the RegionGauges Screen (refer back to Figure 2.2) present the proportion of inmates that fall within the green, yellow, or red categories in terms of their predictive probability. COTAS refers to this predictive probability as a risk score (sometimes referred to as an “Inmate Risk Factor”). Thus, an inmate with a risk score of 32.25 has a predicted 32.25% chance that he/she will be involved in a violent event in the next 30 days. Scores are limited to values ranging between 0 and 100. Inmates with a risk score of less than 30 are coded as “green,” inmates with a score between 30 and less than 80 are coded as “yellow,” and inmates with a risk score of 80 or above are coded as red. For example, an inmate with a risk score of 60 has a predictive chance of involvement in a violent event of 60% and would be coded as “yellow.”

The RegionPredictorDashboard (Figure 2.6) can be navigated to by either clicking on one of the regional “lifesavers” or clicking on the region title above the “lifesavers” on the RegionGauges dashboard. The RegionPredictorDashboard displays two “lifesavers.” The Inmate Predictor on the left shows the proportion of inmates in each color category and is identical to the “lifesaver” on the RegionGauges dashboard. The Facility Predictor “lifesaver” on the right displays the proportion of facilities color-coded according to the probability that the predicted number of inmates who are involved in violent events will be reached within the next 30 days.

Figure 2.6: RegionPredictorDashboard Screen

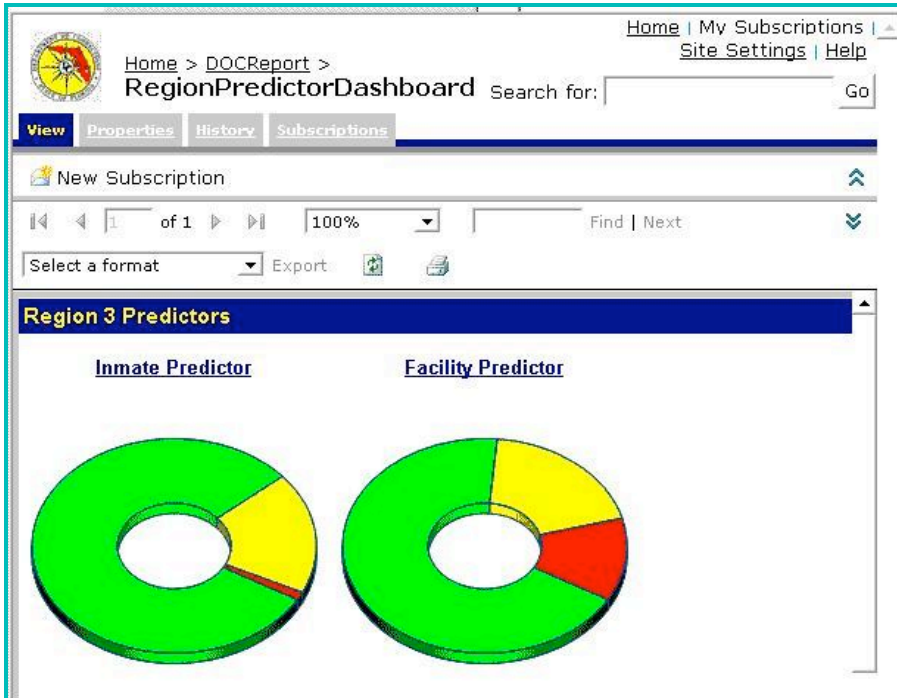


Figure 2.7 presents an example of the RegionPredictorList screen. Users may navigate to the RegionPredictorList screen by either clicking on the Inmate Predictor title or the “lifesaver” in the RegionPredictorDashboard screen. The RegionPredictorList screen allows users to view events at the facility level and examine the proportion of inmates at each facility within the different color categories. All of the facilities in the region are listed in order of the average inmate risk score. For example, the facilities with a high number of inmates who have high risk scores are at the top of the list. The bars under the “risk” column display the proportion of inmates in each risk category. Clicking on the name of the facility (in the left column) will provide the user with a list of each inmate at the facility along with the risk score (see Figure 2.8: PedictorSummaryByDorm Screen).

Figure 2.7: RegionPredictorList Screen

The screenshot shows a web application interface for 'RegionPredictorList'. At the top, there is a navigation bar with links for 'Home', 'My Subscriptions', 'Site Settings', and 'Help'. Below this is a search bar and a breadcrumb trail: 'Home > DOCReport > RegionPredictorList'. The main content area features a table with three columns: 'Facility', 'Facility Type', and 'Risk'. The 'Risk' column contains horizontal bars with segments in green, yellow, and red, representing different risk levels. The table lists 18 facilities, including various C.I. (County Jail) and W.R.C. (Work Release Center) locations.

| Facility | Facility Type | Risk |
|--------------------------------------|---------------------|--------------------------------|
| BREVARD C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| SUMTER B.T.U. | BOOT CAMP | High Risk (Green, Yellow, Red) |
| CFRC-MAIN | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| ZEPHYRHILLS C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| LAKE C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| TOMOKA C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| DEMILLY C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| AVON PARK C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| MARION C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| PUTNAM C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| CFRC-EAST | MAJOR INST. ANNEX | High Risk (Green, Yellow, Red) |
| SUMTER C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| POLK C.I. | MAJOR INSTITUTION | High Risk (Green, Yellow, Red) |
| BARTOW W.R.C. | WORK RELEASE CENTER | High Risk (Green, Yellow, Red) |
| COCOA W.R.C. | WORK RELEASE CENTER | High Risk (Green, Yellow, Red) |
| KISSIMMEE W.R.C. | WORK RELEASE CENTER | High Risk (Green, Yellow, Red) |
| REENTRY CTR OF OCALA | CONTRACTED W/R FAC | High Risk (Green, Yellow, Red) |
| DAYTONA W.R.C. | WORK RELEASE CENTER | High Risk (Green, Yellow, Red) |

Depicted in Figure 2.8, is the PedictorSummaryByDorm screen which allows users to “drill down” to the inmate level of analysis and view each inmate’s name, location, primary work assignment, and risk score. Placing the cursor over an inmate’s risk score displays information about the inmate’s prior violent events and institutional history (see the pale yellow box in Figure 2.8). Tabs at the top of the screen allow users to sort by inmates’ bed assignment, bed mission, facility dormitory, or primary work assignment, thus allowing for semi-relational database functionality. For example, users can identify when inmates with high risk scores are concentrated together. Orange risk scores represent inmates that have scores greater than 50 and have been in the current facility less than six months. Users can obtain detailed information regarding individual inmates from the Correctional Offender Information Network by clicking on inmates’ DC Number and users can obtain information regarding inmates’ involvement in violent events by clicking on inmates’ risk scores.

Figure 2.8: PredictorSummaryByDorm Screen

Home > DOCReport > PredictorSummaryByDorm

Search for: Go

View Properties History Subscriptions

New Subscription

100% Find | Next Select a format Export

Inmate Risk Assessment for Facility: BREVARD C.I. Total Inmates: 992

| Bed | Bed Mission | DC Number | Name | Dorm | Primary Work Assignment | Risk |
|--------|-----------------|-----------|-----------------------|------|----------------------------|-------|
| D3204U | CHILD NUTRITION | B08168 | SALAS, RAUL | D | WELLNESS EDUCATION PROGRAM | 87.89 |
| B3104U | CHILD NUTRITION | B08286 | TORRES, WAYNE | B | ACADEMIC STUDENT | 86.82 |
| C4201U | CHILD NUTRITION | B08708 | ROSALLES, OVIDIO | C | AUTO TECH/AUTOTRONICS | 86.74 |
| D1203U | CHILD NUTRITION | L85708 | LEBRECHT, JESSE | D | WELLNESS EDUCATION PROGRAM | 86.74 |
| H4102L | CHILD NUTRITION | E38822 | CARTER, CODY | H | ORDERLY-EDUCATION | 86.74 |
| C1105L | CHILD NUTRITION | C04638 | KANE, DYLAN | C | WELLNESS EDUCATION PROGRAM | 86.74 |
| E3205L | CHILD NUTRITION | I11540 | RODRIGUES, ROBERTO | E | WELLNESS EDUCATION PROGRAM | 86.74 |
| D2102U | CHILD NUTRITION | T66183 | PEARSON, TAYLOR | D | TRANSITION PROGRAM | 86.74 |
| H3101U | CHILD NUTRITION | R51944 | BYRNE, JOHN | H | ACADEMIC STUDENT | 86.73 |
| C2101U | CHILD NUTRITION | R67666 | RIKARD, COREY | C | WELLNESS EDUCATION PROGRAM | 86.73 |
| E2103L | CHILD NUTRITION | H34620 | VARGAS, EMILIO | E | WELLNESS EDUCATION PROGRAM | 86.72 |
| B3202L | CHILD NUTRITION | R57052 | SCULLION, CHRISTOPHER | B | RESIDENTIAL CARPENTRY | 86.27 |
| B1202L | CHILD NUTRITION | Y41010 | CARDONA, WILLIAM | B | HOUSEMAN | 86.15 |
| C4204L | CHILD NUTRITION | B07464 | PEDRO-MENDEZ, PETER | C | WELLNESS EDUCATION PROGRAM | 86.14 |

Tooltip for D3204U:
 Prior Violent Events: 4
 Prior Non Violent Events: 2
 Time Served: 1 years
 8 month(s) since last violent event
 Violent Offender Flag
 Transfer In Date: 4/23/2010

The RegionPredictorList screen (Figure 2.9) provides the color-coding of facilities based on the predicted probability that a certain number of inmates involved in violent events will reside within a particular facility within the next month. The RegionPredictorList can be navigated by clicking on either the Facility Predictor title or the “lifesaver” on the RegionPredictorDashboard screen (refer back to Figure 2.6). The RegionPredictorList remains under development and offers limited utility in identifying predictors of violent events at the facility level. Indeed, only one independent variable is included in the model: different gang rate (the number of different gangs in a facility divided by the number of inmates in a facility). Similar to the aforementioned color-coded designations, it is unclear how the values assigned to the color-coded categories based on predicted probabilities were established and how the predictive accuracy of the algorithm was assessed.

Discussion

This chapter provides an overview of the information available to users via the COTAS web-based interface. Statistics are presented throughout a number of screens which allow users to explore and view data at various levels (e.g., regional, institutional, inmate). Similarly, inmate characteristics, such as risk scores and involvement in violent events, can be viewed in aggregate at the dorm, facility, and regional levels of analysis. When examining events within a facility, the COTAS interface provides a detailed account of the violent and non-violent events that occurred within the last 30 days. Additionally, the COTAS system provides a 12-month trend of each category of events for the facility and generates a line graph of changes in the number of events over time. When examining events over the past 30 days, COTAS provides detailed information about the event and inmates involved. Details of events are provided with direct links to DOC’s Disciplinary report/Investigation database. Detailed information about inmates who were involved in violent events is provided by links to the Corrections Offender Information Network’s Inmate Population Information Detail. These features make

COTAS a useful tool in providing a general overview of events at a facility and providing users the option to explore each event in great detail.

Chapter 3

Validation of the COTAS Predictive Models

Introduction

The goal of COTAS is to reduce the number of violent incidents in facilities by providing prison administrators with timely and accurate information regarding the risk of violence within their facilities. In order to accomplish this goal, COTAS must first identify the facility and inmate characteristics that are predictive of violence and identify the inmates who are most likely to be disruptive. Indeed, COTAS' ability to predict the likelihood of future violence allows administrators to take actions that *may* reduce the occurrence of violent events. This chapter provides describes the examination of the appropriateness and accuracy of the statistical models or algorithms employed by COTAS. The chapter provides a description of the model used by COTAS as well as the variables used to make predictions. Additionally, the chapter discusses some of the concerns with the modeling procedure and the effect on the overall predictive accuracy of the models. The methods used to validate COTAS's predictions are described and the findings from the validation are presented. The chapter concludes with a summary discussion of the modeling techniques and level of accuracy of COTAS' predictive modeling.

Model description

The development of the predictive models used by COTAS involved a multi-stage process across two DOC offices. BRDA identified the predictors of violent events from DOC's collection of databases, which provide detailed information on the characteristics of individual inmates, violent and non-violent incidents, and environmental characteristics of facilities. OIT and its contractor used the list of predictors from BRDA to create a predictive model or algorithm that would make predictions based on real-time data.

An initial concern with developing the model across two offices is that the two parties are using different software and statistical techniques to identify key predictors of violent events. BRDA used five years of monthly snapshot data to determine statistically significant predictors of inmates' involvement in violent events. Analysis was conducted using SAS (a statistical software package commonly used in business and science). The final list of predictors was chosen based on the predictive accuracy of the measure, the malleability of the measure (i.e., could facility administrators take actionable steps to mitigate the odds of an event occurring), and the parsimony of the overall model. Additionally, some of the BRDA-recommended measures were omitted for unknown reasons and some measures were excluded due to multicollinearity, as discussed in Chapter 1.

The contractor, *Idea*, developed the COTAS predictive algorithms using 2005 Analysis Services in Microsoft Data Mining software, a software package that integrates with Microsoft SQL Server to provide business intelligence services to large relational databases. Microsoft Logistic Regression Algorithm was selected because the predicted outcome was bivariate (violent event/ no violent event) and the distribution of the data did not violate any of the assumptions of logistic regression. The Microsoft Logistic Regression Algorithm requires users to specify two data sources: the training data-set and the live data-set. The training data are historical data used to create the algorithm and the live data is the real-time data for which predictions are applied. For example, the live data in COTAS is the data from DOC's mainframe databases that have been transferred to the data warehouse and is updated nightly. Microsoft Data Mining runs logistic regression models Monday through Thursday evenings using data from the training data-set to create an algorithm for calculating the predicted probability of violence at the inmate level. The training data may be either dynamic (updating with new data) or static, remaining constant. In COTAS, the training data are static and include 1,407,577 inmate records from June 1, 2009 to May 30, 2010. *However, according to the COTAS Technical Documentation from Idea, the training data-set was designed to be dynamic.*

By default, the Microsoft Logistic Regression Algorithm identifies predictor variables to be included in the algorithm by the predictive power of the variable (i.e., the strength of association or correlation between the predictor variable and the outcome variable) holding constant other predictors. The Microsoft Logistic Regression Algorithm "recommends" variables to include and exclude from the model.

The staff from BRDA and the contractors from *Idea* identified predictors of future violence using different data sources, statistical packages, statistical techniques, and selection criteria. This may, in part, explain differences found between the key predictor variables identified by staff from BRDA versus predictor variables identified by *Idea*. In conversations with a consultant from *Idea*, John Marsh, and BRDA staff, it was revealed that there was an ongoing discussion regarding the predictor variables to include in the COTAS algorithms throughout the development of COTAS; however, a consensus as to the most appropriate and significant set of predictors to include in COTAS was not attained. Thus, COTAS currently includes only the predictors identified by *Idea* as significant.

Predictive Variables

The variables included in the inmate level logistic regression are:

- involvement in a violent event (DV),
- inmates' age,
- gender (sex),
- bed category,
- number of months since an inmate's last violent event,
- number of prior violent events,
- positive drug test,

- race (white/non-white)²,
- time served,
- violent offender flag, and
- prior placement in a closed management bed category.

1

There is limited documentation provided by Idea regarding the coding of these variables and the “point and click” interface of Microsoft Logistic Regression Algorithm does not allow for a review of the method for entering variables into the models. Thus, it is unclear what diagnostic tests were conducted to determine the final models. Additionally, a *codebook or data dictionary should have been included* in the COTAS documentation. This document would include a list and descriptions of variables in the models, all of the possible values for categorical variables, and the value labels for each categorical variable.

As discussed in Chapter 2, COTAS is also running an algorithm that predicts the probability of a set number of inmates involved in violent events at the facility level. *Currently, there is no documentation on this predictive model; however, it is clear that only one independent variable is included in the model: different gang rate.* The different gang rate is the number of different gangs in a facility divided by the inmate population of the facility.

Causal ordering and statistical power

Another factor that contributes to the discrepancy between the predictive models generated by BRDA and those generated by the *Idea* contractors is the structure of the data warehouse. *When new events or changes in inmate status are entered into the data warehouse, all data changes are automatically assigned as occurring on the first day of the month.* For example, if an inmate is involved in a violent event on June 27th, it is recorded in the data warehouse as occurring on June 1st. Thus, the COTAS algorithm does not predict the probabilities of inmates’ involvement in violent events over the following 30 days, but instead predicts the probability of inmates’ involvement in a violent event anytime during the month (lacking specific time order). For example, inmates’ risk scores for June predict the probability that they will be involved in a violent event in June rather than predicting that an inmate will be involved in a violent event during the next 30 days. Having everything entered on the same date (the first day of the month) *creates a significant problem for the predictive models* because the models can not consider the causal ordering of any events in a particular month. For example, it may appear that inmates in disciplinary confinement are involved in higher rates of violent events, when the reverse is most likely occurring (violent events result in inmates being placed in disciplinary confinement).

² DOC may want to consider removing or revising the race variable (operationalized as white/all other races) used in COTAS because it oversimplifies the effect of race (i.e., being white is most likely a proxy measure of some other construct) and because there is no clear explanation regarding why whites should be singled out for the benefit of lower inmate risk scores compared to other racial/ethnic groups.

Indeed, because the data warehouse updates records nightly, the causal ordering issue *becomes a significant problem, first because the model is incorrectly specified, but more importantly because it impacts the accuracy of all of the predictor variables in the model.* For example, building on the scenario above, if the causal ordering is reversed and disciplinary confinements are used to predict inmates' involvement in violent events, then we would expect the two variables (disciplinary confinements and violent events) to be strongly correlated. This strong correlation in the model leaves all the other variables with a smaller amount of unexplained variance to predict. This has the effect of artificially diminishing the influential power of other predictors in the model such as age, sex, or number of prior violent events. This may have resulted in key variables being excluded from the model when, in fact, they were significant predictors of violence. *There seems to be evidence that there are significant causal ordering problems with the COTAS algorithms.* First, the contractors from *Idea* identified fewer predictive variables as significant relative to the variables indicated by BRDA—despite the fact that BRDA had more stringent criteria for inclusion into the model. Second, as will be discussed in greater detail in this chapter, the COTAS algorithms are over-estimating the probability that an inmate will be involved in a violent event.

One final issue related to the impact that changing the date for all events to the first day of the month has on the predictive accuracy of the COTAS algorithms is that inmates' data from Fridays and Saturdays are not recorded until Sunday night, which introduces additional error into the COTAS predictive models if the first day of the month occurs over a weekend. For example, May 1, 2011 occurred on a Sunday. Thus, events that occurred on April 30th (Friday) and April 31st (Saturday) were not recorded in COTAS because they occurred in April. This is a relatively minor issue, but it introduces additional stochastic error, thus further diminishing the predictive accuracy of the COTAS algorithms.

Overall, the statistical software (2005 Analysis Services in Microsoft Data Mining software) and the statistical technique (logistic regression) are appropriate choices for the predictive modeling in COTAS. COTAS could improve the predictive accuracy of the models by: (1) changing the training data-set from static to active (i.e., the previous 12 months of data), (2) addressing the causal ordering issue by using the actual dates of events and, thereby, having the Microsoft Logistic Regression Algorithm model the predicted probability of inmates' involvement in violent events *over the following 30 days*, and (3) updating the existing algorithm to reflect these changes. Many of the errors in the statistical modeling and the creation of the COTAS algorithms discussed above may have been avoided if there had been greater communication between contractors from *Idea*, BRDA, and OIT. Documentation from *Idea* should have clearly indicated that *the COTAS algorithms were predicting inmates' involvement in violent events using non-time ordered variables.* Additionally, the inclusion of a codebook or data dictionary may have allowed reviewers to easily identify some of the errors in the modeling procedures during a pilot test.

Validation Methodology

The validation of the predictive accuracy of COTAS was conducted by comparing the *predictive probability* that an inmate will be involved in a violent event within 30 days

with the inmate's *actual involvement* in a violent event over a 30-day period. Project staff made the determination to examine 30 days of data from COTAS in order to overcome the problem created by the fact that COTAS changes the date of all events to the first day of the month. Therefore, project staff examined risk scores prior to the occurrence of a violent event and excluded risk scores that occurred after an event has taken place *which avoids some of the causal ordering problems*.

Using data warehouse files (cos_19 Dim_InmatePredictions and StagingdmInmatePrediction), data on inmates' risk scores (probability) and involvement in violent events (event_count) were extracted daily from May 1st to May 30th2011. Inmate records were merged first using the offender_id to merge inmates' risk scores and involvement in violent events within the same day, and then the variable bkdcnumber from the dim_offender file was used to merge inmate records together from multiple days. The merged data-set includes 110,030 inmates. Inmates with missing data were excluded from the analysis resulting in a total of 101,485 inmate records. Analysis of missing data indicated that inmates with at least one or more missing data values had higher risk scores and lower involvement in violent events relative to other inmates; however, these differences were relatively small and would not change the overall findings of this validation.

The measure of inmate involvement in a violent event was coded as 0 = inmate was not involved in a violent event during the 30 days of observation, 1 = inmate was involved in a violent event during the 30 days of observation. The predicted probability of an inmate's involvement in a violent event over a 30-day period was calculated as the sum of the daily predicted probability of an inmate's involvement in a violent event divided by the number of days until an involvement in a violent event occurs. If an inmate was involved in no violent events during the 30 days of observation, then the number of days until an involvement in a violent event has occurred equaled 30.

In the following section, findings from the COTAS validation are presented in three steps. First, the predictive accuracy of the COTAS algorithm is examined in terms of the predetermined color-coding presented in the user interface (green, yellow, and red). The focus of this analysis is to determine if the color-codes make meaningful distinctions between offenders in terms of their involvement in violent events. Second, an examination of the percentage of inmates involved in violent events by each deciles division is presented. This analysis examines the discriminatory power of the risk score. Additionally, the overall predictive power of the risk score is discussed. Third, the sensitivity and specificity of the COTAS predictions are examined whereby, inmates given a predicted probability greater than or equal to .5 (a risk score of 50) are predicted to be involved in violent events and inmates with a predicted probability of less than .5 are predicted to not be involved in violent events.

Findings

Of the 101,485 valid inmate records collected from May 1, 2011 to May 30, 2011, roughly 1% of inmates (987) were involved in one or more violent events. However, the average inmate risk score among the 101,485 inmates was 29.7, suggesting that the COTAS Algorithm projected that roughly 30% of inmates (30,445) would be involved in violent events. Table 3.1 indicates that the average risk score for inmates involved in a

violent event was 42.5 whereas the average risk score for an inmate who did not engage in a violent event was 29.6. Thus, the risk scores for inmates involved in violent events were slightly higher than inmates not involved in violent events.

Table 3.1: Inmate Risk Score by Involvement in Violent Events

| | Average Inmate Risk Score | Median Inmate Risk Score |
|--------------------------------|---------------------------|--------------------------|
| Not Involved in Violent Events | 29.6 | 30.1 |
| Involved in Violent Events | 42.5 | 42.3 |

Table 3.2 indicates the average color-code of the inmates in the sample during the observation period. Most inmates were coded as green (risk score of less than 30) or yellow (risk score ≥ 30 but < 80), with over half of the sample coded as yellow. Only 193 inmates or 0.2 percent were coded as red (risk score ≥ 80).

Table 3.2: Descriptive Statistics of Inmates across the Color-Coding Categories of Risk Levels Displayed by COTAS

| Color-Code | N | Percent | Average Inmate Risk Score |
|------------|--------|---------|---------------------------|
| Green | 50,344 | 49.6% | 17.8 |
| Yellow | 50,948 | 50.2% | 41.3 |
| Red | 193 | 0.2% | 82.3 |

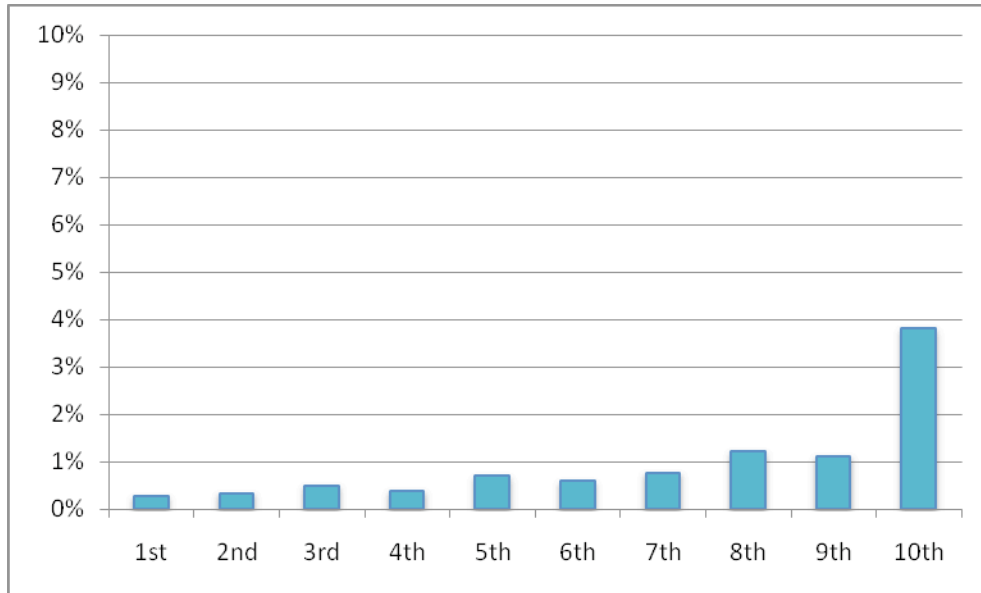
Analysis presented in Table 3.3 examines the color-coding scheme's level of accuracy in assessing inmates' involvement in violent events. Results indicate that inmates who were coded as yellow were involved in more violent events, both proportionately and numerically, relative to inmates coded as green or red. However, there is no statistical difference between the proportion of inmates involved in violent events in the red category versus the proportion in the yellow or green. *In other words, inmates in the red category are statistically indistinguishable from inmates in the green or yellow categories in terms of their involvement in violent events.* Additionally, inmates in the yellow category were significantly ($p < .001$) more likely to be involved in violent events relative to inmates in the green category. Thus, while the color-coding out performs what we would expect to see by *random chance*, these results suggest that the existing COTAS *algorithm is inefficient*. Specifically, no statistically significant differences in inmates' involvement in violent events were found between the red category and yellow category, and no differences were found between the red category and green category. Differences in inmates' involvement in violent events were found only in the green category and the yellow category.

Table 3.3: Inmates Color-Code by Involvement in Violent Events

| Color-Code | Percent Involved in Violent Events | Inmates Involved in Violent Events |
|-------------------|---|---|
| Green | .4% | 216 |
| Yellow | 1.5% | 769 |
| Red | 1.0% | 2 |
| Total | 1.0% | 987 |

Figure 3.1 displays the percentage of inmates involved in violent events in each deciles group of risk scores. In other words, impose a ranking system on the inmates’ risk scores from highest to lowest resulting in the scores in the bottom 10 percent would be in the 1st deciles group, scores from the 10th to 20th percentages would be in the 2nd deciles group, etc. Each deciles group represents roughly 10,149 inmates (10 percent of entire sample). Of inmates with risk scores that were at the bottom 10 percentile, .28 percent were involved in violent events whereas 3.81 percent of inmates in the top 90th percentile of risk scores were involved in violent events. Further, 39 percent of all inmates involved in violent events were inmates with risk scores that were in the highest 10 percent of risk scores. Thus, the *predictive accuracy of the risk scores seems limited to the highest scoring inmates.*

Figure 3.1: Percentage of Inmates Involved in Violent Events by Deciles Group of Rank Order



The *sensitivity* of a prediction model refers to the proportion of actual violent events that were correctly identified or the ability of a prediction to correctly identify positive results. *Specificity* refers to the ability of a prediction to correctly identify negative results. Keep in mind that a relatively small percentage of inmates are actually involved in violent events (likelihood is relatively rare: approximately 1% of inmates per month). To evaluate the *sensitivity and specificity* of the COTAS algorithm, FSU converted risk scores into specific predictions, whereby inmates whose risk score is greater than or equal to 50 are predicted to be involved in a violent event and inmates with risk scores of less than .5 are predicted to abstain from violent events³. Table 3.4 displays the predictions of inmates’ involvement in violence using the above recode of risk scores.

The COTAS algorithm correctly predicted inmates’ involvement in violent events 334 times (sensitivity); this value is referred to as a “true positive.” “True negative” findings occurred 92,862 times, whereby the COTAS prediction that a violent event would not occur was correct. “False negatives” (Type II error) occurred 653 times and “false positives” (Type I error) occurred 7,636 times. In other words, COTAS *wrongly predicted* that an inmate would be involved in a violent event 7,636 times and COTAS *wrongly predicted* that an inmate would not be involved in violence 653 times. Both types of error are problematic; however, which error is more problematic is a subjective issue, not a statistical determination. Another way to discuss the accuracy of COTAS would be to say that when the COTAS Algorithm predicted involvement in a violent event, it was correct 4 percent of the time, however when it predicted that an event would not take place it was correct over 99 percent of the time.

The specificity of the COTAS model is relatively high (.93). This is not surprising, given that inmates’ involvement in violent events is relatively infrequent. A prediction model with high specificity will have a low rate of false positives and a prediction model with high sensitivity will have a low rate of false negatives. The sensitivity of the COTAS model is .34, indicating that only 34% of the inmates involved in violent events was correctly identified by COTAS. A predictive model based on random chance would have a specificity value of .99 and a sensitivity value of .01.

Table 3.4: Predictions of Inmates’ Involvement in Violent Events by Actual Inmate Involvement in Violent Events

| | <u>Actual</u> | |
|-----------------------------------|---------------|------------------|
| | Violent Event | No Violent Event |
| <u>Predicted</u> Violent Event | 334 | 7,636 |
| No Violent Event | 653 | 92,862 |

³ This calculation was employed because .5 represents the tipping point between the probability that event *will not occur* and the probability that an event *will occur*.

FSU assessed the predictive accuracy of the COTAS algorithms using a sample of 30 days of inmate data. Results indicate that the COTAS algorithm *performs better than random chance*. For example, inmates color-coded as yellow were involved in violent events at a higher rate than inmates coded as green. However, the *predicted probabilities generated by the COTAS algorithms are disproportionately high relative to the actual occurrences of inmates' involvement in violent events*. COTAS estimates that 30 percent of inmates will be involved in a violent incident over a 30-day period, when roughly 1 percent of inmates are actually involved in violent events. Without documentation of the COTAS modeling procedures and coefficients used in the model, it is difficult to determine the source of this error. However, given the relative size of the overestimate (over 30 times the size of the actual event's occurrence), *there are significant errors either in the predictive model or the transformation of values to predictive probabilities*. *Again, without documentation of the modeling procedure, the source of key problems cannot be identified.*

Summary and Discussion

The development of the predictive models for COTAS was a multi-stage process split across two DOC offices: BRDA and OIT. Ultimately, the contractor, *Idea*, created the predictive models used by COTAS; however, it is unclear how much consultation with BRDA occurred. In discussions with BRDA and one of the consultants from *Idea*, it appears that this consultation was limited. No evidence was found to indicate that BRDA was directly included or involved in the creation of the COTAS algorithm once key predictors were identified⁴. Working in separate software packages and with separate data-sets limited the ability of BRDA to oversee the accuracy of *Idea's* statistical models.

The variables in the COTAS algorithms were determined, in part, by feedback from the Microsoft Logistic Regression Algorithm, which indicated the variables that were statistically significant at a predetermined confidence level *which was not documented*. Because the dates of the model were changed to make all events occur on a single day of the month, the model reflects only the month that data were collected (i.e., the first day of the month) and the model predicts only events within that same month; therefore, the *ability to determine the causal ordering of events was eliminated*. *This has a significant impact on the accuracy of the predictive model*, both in terms of which variables were deemed as statistically significant (thus determining which variables were included in the model) and the predictive accuracy of the model.

Findings from the analysis of the predictive accuracy of the COTAS algorithms indicate that the model outperformed random chance alone; however, the color-coding and risk scores had limited utility in correctly identifying inmates that were involved in violent events. Inmates in the red category were statistically similar to other inmates in terms of their involvement in violent events. Inmates in the yellow category were statistically more likely to be involved in violent events relative to inmates in the green category.

⁴ While it became apparent that there was a disconnect between the vision and effort of BR&DA and that of *Idea*; the reason for this disconnect could not be identified.

Chapter 4

The Utility and Functionality of COTAS from the Perspective of the Users

Introduction

Another component of FSU's validation of COTAS is the assessment of the utility and ease of use of COTAS for end-users. To accomplish this task, a user survey was developed to gain an understanding of the extent to which correctional administrators and staff use COTAS and, more specifically, which components of COTAS are most useful. Previously, DOC had developed and administered a pilot survey with the similar intentions. However, the survey results were inconclusive because the survey instrument was limited in scope and the response rate was very low (n=2).

COTAS Research

The first step in developing the survey instrument involved an examination of each COTAS screen to understand its purpose and functions, and to identify issues with the interface that could be useful for developing the survey instrument. FSU reviewed two user's manuals, both prepared by *Idea*. Only the first user's manual was disseminated to administrators in the field; therefore, users' knowledge of COTAS and level of expertise may be, in part, guided by this manual. Further, the survey question that inquires about the user's manual is referring to this first manual. DOC contracted with *Idea* to revise and update the user's manual; however, the second manual had not been disseminated to the field at the time of this validation. The second manual is more comprehensive than the first manual but, to our knowledge, it has not been field tested by users. Next, FSU reviewed the results from the two responses of DOC's pilot survey. The limited results indicate that for the most part the respondents thought COTAS was helpful but respondents did indicate that some risk factors were inaccurate and that they did not completely understand the method by which COTAS categorized certain institutions. This information identified several key areas of COTAS that were addressed in FSU's survey instrument.

Development of the Survey Instrument

From the research of COTAS as well as the pilot survey responses, project staff identified the following areas of inquiry for the survey instrument:

- Title/position of respondent: DOC's Project Manager (Gail Denson) provided a list of the correctional staff and administrator positions that had access to COTAS (Warden, Assistant Warden, Colonel, Major, Captain, and Classification Supervisor). All six positions and an "other (please specify)" category were included in the survey. This question was included because it would be helpful to know if there is variation in the experiences for specific types of correctional administrators (in terms of general usage as well as more detailed information such as screens utilized).

- **Correctional Facility:** DOC provided a list of correctional facilities and their number; respondents were asked to indicate the facility in which they worked. This question was included because it would be useful to know if COTAS use varied across specific types of institutions or regions.
- **COTAS use:** Respondents were asked whether or not they used COTAS. Yes and no responses were directed to different follow-up questions.
- **Prevalence and Frequency of COTAS use:** Respondents were asked to approximate the length of time that they had used COTAS (number of times per day, per week).
- **Use of COTAS functions:** Respondents were asked to indicate their primary reasons for using COTAS (they could select more than one reason). Answer choices included: monitoring events in your facility, monitoring events in your region, monitoring predictions of violence in your facility, monitoring the predictions of violence in your region, inmate search, D.R. lookup, high profile inmates. An “other” option was included along with a please specify text box that allowed them to type in their own responses.

The survey instrument included a section of questions to inquire about the specific screens that respondents used, as well as the frequency with which they use them. Each survey question in this section included the name of the screen (ex. RegionGauges, RegionGraph, RegionPredictorList...ect.), a short description of specifically what each screen does (ex. RegionPredictorList: Facility’s violence risk based on percentage of high risk inmates-click on inmate predictor in the RegionPredictorDashboard or EventGraph: graph of disciplinary reports, investigations, and use of force-violent and non-violent), as well as a screen shot depicting screen addressed by the question (to avoid confusion). For each COTAS screen addressed, responses choices were: never, sometimes, or often. All fifteen COTAS screens were addressed to solicit an accurate account of the frequency with which respondents used specific screens. An optional question was included asking the respondents to indicate other functions of COTAS that they used (and the frequency).

NOTE: Response categories:

For questions asking respondents to rate the usefulness of specific screens, response categories were displayed as a five-point scale: 1 (completely) useless through 5 (very useful).

For questions asking respondents to indicate how often a screen’s information was inaccurate, response categories were: always, sometimes, never, do not know, and other (please specify).

For questions asking respondents how often they used a specific screen, response categories were: never, sometimes, often.

- **Historical screens (monitoring your facility):** Respondents were asked to rate the region facility list (screen entitled RegionDashboard) for accuracy. To avoid confusion and increase the reliability of responses, each question was accompanied by a screen shot depicting the screen in question. The following sentence was included to clarify the scope of the question: “For example, are the facilities located in the proper categories (red, green, and yellow)?” The answer choices were: always, accurate, sometimes accurate, never accurate, do not know and optional (please specify) with a text box. Respondents were also asked to rate the usefulness of the region facility list (screen titled RegionDashboard). The last question for the RegionDashboard screen was: “Were you aware that facilities are ranked depending on category (work release centers, work camps and jails, female institutions...etc.)? This was followed by yes and no answer choices. This question was included to gain insight into why COTAS users do or do not understand (or assuming the system is inaccurate) the categorization of the region facility list (as indicated by the pilot surveys).
- **Historical Screens (monitoring your facility) continued:** The next series of questions addressed the FacilityMonitor screen (a screen shot was included in the question stems). The first question addressed any inaccuracies in this screen, the question stem is: How often have you discovered inaccuracies on the screen titled FacilityMonitor? The next question about the FacilityMonitor screen addressed the usefulness of the screen. The survey asked the respondents to rate the usefulness of event counts in their facility (as shown on the FacilityMonitor screen) on a scale of 1-5 (with 1 being completely useless and 5 being very useful). Choices 1-5 were listed for respondents to choose from and, again an “other (please specify)” option was included.
- **Predictive Screens (predicting and mitigating risk):** There were several questions (referring to the facility and inmate predictor screens) asking respondents to indicate the frequency of use, the usefulness, the accuracy of each of the predictive screens, and to indicate their level of understanding of the screen’s information (distinction between red, yellow, and green for the inmate predictor or facility predictor). Each question included a screen shot of the predictive “lifesavers.” If respondents indicated that they did not use a particular screen, the survey directed them to a follow-up question to indicate the reason (no accurate, not helpful, other). If respondents indicated that they understood the distinctions between the color codes, they were asked to rate the level of accuracy of the predictors. If respondents indicated that the predictors were inaccurate, they were asked why (thresholds are arbitrary, predictive variables are incorrect, do not know, all, and other). This series of questions was included to gain an understanding the following:
 - Do the correctional administrators and staff understand the distinctions between the red, yellow, and green for the inmate predictor?
 - If so, do they think the distinctions are accurate?

- Why do you think they are inaccurate (if you answered no to the following question)
- **Final Thoughts:** This final series of questions addressed the overall functionality of COTAS as well as suggestions for improvements from the correctional administrators and staff. Respondents were asked if COTAS aids their ability to prevent the occurrence of violent events; ways that the prediction dashboard could be improved; and would an updated user's manual improve your experience with COTAS. These questions were included because the results from the pilot survey indicated that the prediction dashboard was rarely used because they believed that it was inaccurate. The open-ended response option for the question regarding improvements to COTAS was an attempt to solicit comprehensive, accurate, descriptive answers. The researchers felt that a closed-ended question would not be comprehensive enough to render reliable results.

Administration of the Survey

Prior to disseminating the survey to the correctional administration and staff, the survey was subjected to several internal reviews by FSU and DOC headquarter personnel. DOC provided the project staff with a comprehensive list of all wardens, assistant wardens, colonels, majors, and captains (respondents). Through the help of the OIT, a mailing list was established (COTASInstitutions@mail.dc.fl.us) to facilitate dissemination. The survey was administered using the online survey application SurveyMonkey. At the suggestion of DOC's project manager, FSU research staff agreed for the initial email to be sent from the Assistant Secretary of Institutions, Russell Hosford. This email would alert respondents that FSU was conducting a validation of COTAS and this survey was part of that effort. The email text also encouraged respondents to participate in the survey. The mail, from DOC, included a link directly to the survey on the Internet. There was a discussion among project staff as to the appropriateness of the link to the initiate the survey being included in the DOC email (sent from DOC personnel) or if the survey link should be sent from the FSU project staff. DOC believed it would increase the response rate if the link came from DOC; however, in an attempt to minimize potential bias, the FSU research staff wrote the body of the email sent by Mr. Hosford. The email that Mr. Hosford sent to the respondents (authored by FSU research staff) read:

The Florida Department of Corrections has asked the research staff at Florida State University to gather information about the use of COTAS. We are interested in the extent to which Correctional Administrators and Staff use COTAS and more specifically which components of COTAS. Your participation in this survey is for research purposes only and the information provided will be used to recommend improvements to COTAS. This short survey should take less than ten minutes and your responses will be confidential and anonymous. Furthermore, the data will only be reported in the aggregate, meaning specific answers cannot be linked to the individual who supplied them. Below you will find the link to the survey. Your participation is essential to determine how improvements to COTAS can be made. Thank you for your time and contribution to this important project.

The respondents received this email on Friday May 20th. The following Wednesday (May 25th) the respondents received an email from FSU via surveymonkey.com urging them to complete the survey at the request of FSU research and Assistant Secretary of Institutions Russell Hosford by Friday May 27th.

Results

Response Rate and Demographics

Due to time limitations, respondents were given only one week to complete the survey. Within that limited time frame, 167 completed surveys were received from a sample of approximately 300 respondents who were asked to complete the survey, yielding a response rate of slightly higher than 50 percent. The range of responses were fairly evenly distributed across professional position type/job title, with assistant wardens having the highest concentration (28%), followed by majors (21.4%), classification supervisors (19%), wardens (17.9%), and captains (2.4%). Only 15/168 respondents skipped the question asking for their position/job title. The range of the workforce of Florida's correctional institutions was fairly well represented: out of the 75 parent facilities, at least one staff member from each institution completed the survey with the exception of the following 16 institutions:

| | |
|---------------------------|-----------------------|
| 112 Bay C.F. | 404 Okeechobee C.I. |
| 114 R. Junction Work Camp | 405 South Bay C.I. |
| 121 Liberty Work Camp | 417 Lantana C.I. |
| 201 New River C.I. O-Unit | 462 Glades Work Camp |
| 219 Lake City C.F. | 500 Region 5 Office |
| 277 Gainesville C.I. | 511 Moore Haven C.F. |
| 300 Region 3 Office | 529 Hillsborough C.I. |
| 400 Region 4 Office | 576 Hendry C.I. |

Therefore, at least one survey response was received from 79 percent of the institutions that received the survey.

Out of 182 respondents, 57.7 percent (n=105) used COTAS, while 42.3 percent (n=77) did not use COTAS. Of the 77 respondents who did not use COTAS, 71 provided a reason. The most frequently cited reasons are:

- I did not know I had access to it (n=9)
- Not enough training to navigate it successfully or not familiar with the system (n=28)
- COTAS is not useful to them/get the information elsewhere (n=32)

Due to the SurveyMonkey setting applied by the researchers, a skip pattern was applied to the survey. This allowed the respondents to answer different questions depending on a particular response. Therefore the sample size is not consistent for each question. The

responses discussed in the results section were answered only by those respondents who replied that they did use COTAS, maximum sample size of n=105.

Frequency of COTAS use

With regard to the frequency with which the respondents used COTAS, the average length of time the respondents have been using COTAS was 21 months (n=102), while the mode (or most frequently occurring number) was slightly higher at 24 months. The length of reported use ranges from as brief as 1 month to as long as 72 months (*72 months may be a nonplausible value/response because the first version of COTAS was released only 36 months prior to the survey*). Most respondents indicated they used COTAS once a day or less than daily (n=93). On average, respondents reported that they use COTAS 1.6 times per week, but not more than 5 times a week.

Use of COTAS Functions

Respondents were able to identify multiple reasons for their use of COTAS. Ninety-eight respondents answered this question and the most cited reason is “monitoring events in their facility” (83%, n=81). The next most frequently cited reasons are to monitor the predictions of violence in their facility (58%, n=57) and viewing high profile inmates (55%, n=54). Other frequently cited reasons for using COTAS include:

- to monitor the predictions of violence in their facility (slightly more than 50%, n=50),
- to monitor the predictions of violence in their region (29%, n=29)
- to look up disciplinary referrals (18%, n=18)
- to conduct inmate searches (13%, n=13)

The majority of the respondents used COTAS to monitor the events in their facility, while the least amount of respondents used COTAS for inmate searches.

Regarding the frequency with which respondents used various COTAS screens, generally speaking, respondents used COTAS screens fairly evenly, with the exception of a few screens. In most responses, the respondents indicated that they used each screen “sometimes” (rather than “often” or “never”); however a few screens/functions of COTAS appear to be more frequently and less frequently used:

Least popular (these three screens/functions received the highest percentage of respondents who *never* used the screens):

- DisciplinaryReports_Investigations screen: 43% of respondents reported never using this screen
- Inmate Population Information Detail screen: 46% of respondents reported never using this screen
- Gang detail screen: 40% of respondents reported never using this screen

Most popular: (these three screens/functions received the highest percentage of respondents who frequently use the screens (use “often”)):

- Region Gauges screen (30 day event history for region): 22% of respondents reported using this screen often
- Region Dashboard screen (details about specific region’s facilities): 31% of respondents reported using this screen often
- Facility Monitor screen (30 day event history for specific facility): 43% of respondents reported using this screen often

Historical Screens: Monitoring your Facility

Regarding the accuracy and usefulness of the historical screens for the Region Dashboard, respondents reported this screen to be “sometimes accurate” and “useful.”

Accuracy:

- 49% sometimes accurate
- 39% did not know
- 8% always accurate
- 3% never accurate

Usefulness (1=useless and 5=very useful):

- 46% rated 3 out of 5
- 31% rated 4 out of 5
- 6% rated 5 out of 5

Awareness of facility ranking by category (work release centers, work camps, jails, female institutions):

- 59% Aware
- 41% No aware

Regarding the accuracy and usefulness of the historical screens for the Facility Monitor, respondents reported this screen to be “sometimes accurate or not sure of accuracy” and “useful.”

Accuracy:

- 26% sometimes accurate
 - 39% did not know
 - 35% always accurate
 - 0% never accurate
- (included a comment “I didn’t look at it that closely”)

Usefulness (1=useless and 5=very useful):

- 51% rated 3 out of 5
- 27% rated 4 out of 5
- 8% rated 5 out of 5

Predictive Screens: Predicting and Mitigating Risk

Regarding the accuracy and usefulness of the predictive screens for the Prediction Dashboard (also known as the “lifesavers”) respondents reported this screen to be “sometimes accurate or not sure of accuracy” and “useful.”

Frequency of Use:

60% used sometimes
38% used never
2% used often

Reason for Not Using prediction dashboard (reported by 38% who reported never using this screen):

56% it was not helpful
11% it was not accurate
42% provide alternative reason for not using this screen:

Other reasons include:

- Past history and experience has not required this to be a priority
- Time
- I really do not know how to use it
- The information does not provide a true picture of the risk presented at a facility. For example: a facility that rarely has a use of force or averages just a few each month can quickly be in the red if one inmate decides to be disruptive and there are several uses of force against this one inmate. Although the facility may be in the red, it is not a true indication that the facility is under undue stress.
- not helpful for type of facility, issues are monitored by classification & security as well as administrators when events occur
- Not readily available
- Haven't gotten in to the habit of utilizing COTAS
- Just have not
- Had no need for use
- I'm aware anything could happen at any time to change the atmosphere in an institution
- Unfamiliar on how to access it
- I looked at the screen but I kept my own data and went by my data not the COTAS information.
- Primarily interested in my Institution.
- Haven't used it before
- Was not aware of this capability

Usefulness (1=useless and 5=very useful):

11% rated 1-2 out of 5
56% rated 3 out of 5

27% rated 4 out of 5
5% rated 5 out of 5

Portions of the PredictionDashboard screen (could select both):

86% used Facility Predictor
51% used Inmate Predictor

Understanding the distinctions between colors (red, yellow, green):

85% understood distinctions
15% did not understand distinctions

Accuracy of color distinctions for specific institution:

76% accurate
18% did not know
6% not accurate

Reasons for inaccuracy (could check more than one response) (n=4):

50% thresholds are arbitrary
50% don't know why it is inaccurate
25% predictive variables are inaccurate
25% other: "The system uses all grievances as part of the predictors which in my opinion gives a false or inaccurate reading, when the grievance concerns non-institutional issues."

Regarding the functions of the Prediction Dashboard:

Inmate Predictor:

Usefulness (1=useless and 5=very useful):

49% rated 3 out of 5
33% rated 4 out of 5
8% rated 5 out of 5

Facility Predictor:

Usefulness (1=useless and 5=very useful):

7% rated 1 out of 5
10% rated 2 out of 5
54% rated 3 out of 5
28% rated 4 out of 5
7% rated 5 out of 5

Users' Final Thoughts

Overall usefulness of COTAS ("COTAS aids your ability to prevent the occurrence of violence events"):

9% Strong agreed
70% Agreed

21% Disagreed

Comments:

- Allows monitoring for any trends that could prevent problems and aids in determining if a change is needed in a specific area. I like what it offers and feel that it is a useful tool/resource.
- I do not know about its accuracy with any facility other than mine.
- It can help by reviewing the risk factors of inmates housed in certain dorms. But a lot of the high risk factors are mandated to be housed in certain dorms so therefore is where the most problems are, and seem to come from.
- I have not used it for this purpose since I work in Central Office, not at the institutional level.
- It is useful information. It is difficult to determine when and where violent events will happen.
- if it is utilized by the administration
- This is a history and although mostly accurate, prevention of incidents is unlikely. This tells us what to expect in the form of types of incidents, unfortunately, it does not tell us who the violence will come from.
- COTAS makes you aware of a pattern of events/trends that could assist you in identifying areas of vulnerability; history is the best lesson of probability, thus it does help in prevention.

Suggestions for improvements:

- Explained better, updated more often, and more user friendly.
- Display the data on an overlay of an image of the facility - ie where the inmates with a higher predictive score live and work
- More concise information.
- Risk factors are figured to my knowledge on individual information, but is sometimes influenced by additional gang members housed at certain locations increasing risk factors
- Give access to the OIC's.
- It is good
- Additional breakdown by facility type. indicating work release center in red over FSP because of one walk off escape is not a good indicator or predictor of future events
- I don't know of any ways the system could be improved. It is a useful tool.
- find a way to accurately capture real time data
- I am not sure how it can be approved. This has so many denominators to consider.
- Remove the grievance gauges.
- Safety for staff and Inmates.
- a updated summary on each institution, (how many STG members, violent offenders, escape risk assaults on staff and inmates, how many inmates are

requesting PM, ETC.) this would keep you from having to pull up so many different screens

Update User's Manual⁵:

90% reported yes, their experience would be enhanced by an updated manual

10% reported no, their experience would not be enhanced by an updated manual

Comments:

- Only when something new is added
- Received very little training after initial start up. Additional training would be beneficial.

Filtered Responses

The online survey program, SurveyMonkey, allows the researcher to filter the responses, holding one question constant. By holding a particular question constant, the effect on other questions can be isolated. There were two questions in particular that had the potential to greatly affect the respondents' use of COTAS. Both questions related to an understanding of the way that COTAS categorized institutions and made distinctions between the categorizations. DOC's survey responses indicated some confusion for one or both of these aspects.

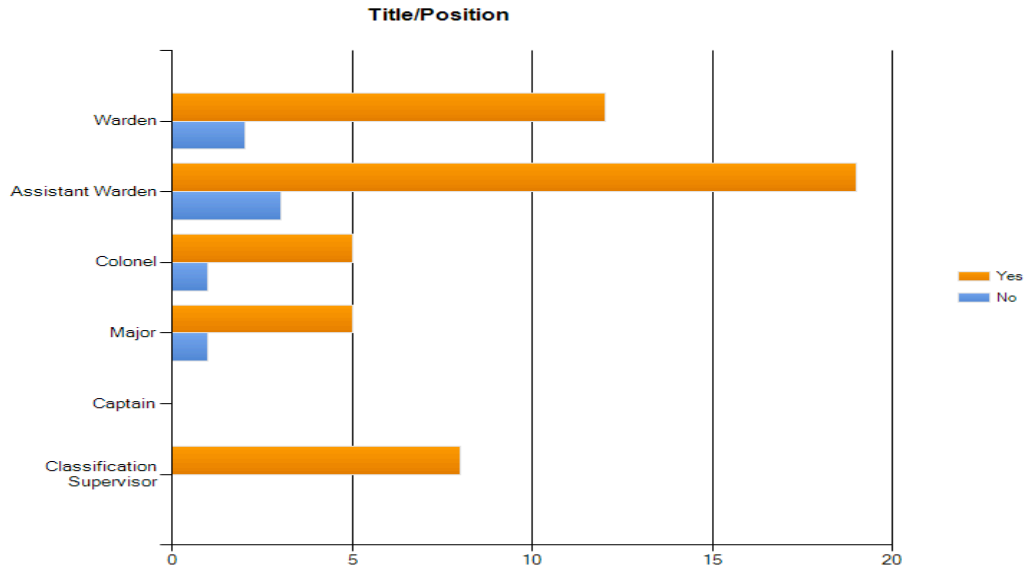
The questions that were held constant are presented in italics at the beginning of each discussion section.

Do you understand the distinction between the red, yellow, and green in the inmate predictor?

Responses from this question indicated that this issue is a significant factor in the COTAS user experience in terms of the functions DOC personnel used and their perception of COTAS' usefulness and accuracy. The color coding refers to the Inmate Predictor "lifesaver" in the Prediction Dashboard entitled RegionDashboard. This question was included in the survey because there was confusion surrounding the understanding of this function of COTAS (demonstrated in DOC's pilot COTAS survey and in FSU's pilot test/review). Generally, there was no discernable pattern between the type of position/title and their understanding of the distinctions between the inmate predictor. Despite this, there was a higher frequency of those who did not understand the distinctions among the wardens and assistant wardens, as opposed to colonels, majors, captains, and classification supervisors.

⁵ This question refers to the first user's manual. As stated previously, while DOC contracted for a revised manual; it had not been tested by or disseminated to the field.

Figure 1: Respondents' Understanding of Color Distinctions across Position/Job Title



Where this question really made a difference was in the use of various COTAS functions and the frequency of using various functions. The respondents who indicated they understood the distinctions were more likely to use a greater number of screens (and therefore more functions). Further, the respondents who understood the distinctions were more likely to use the prediction screens and use them with greater frequency.

Figure 2: Use of COTAS Screens/Functions for Respondents who Understand the Color Distinctions

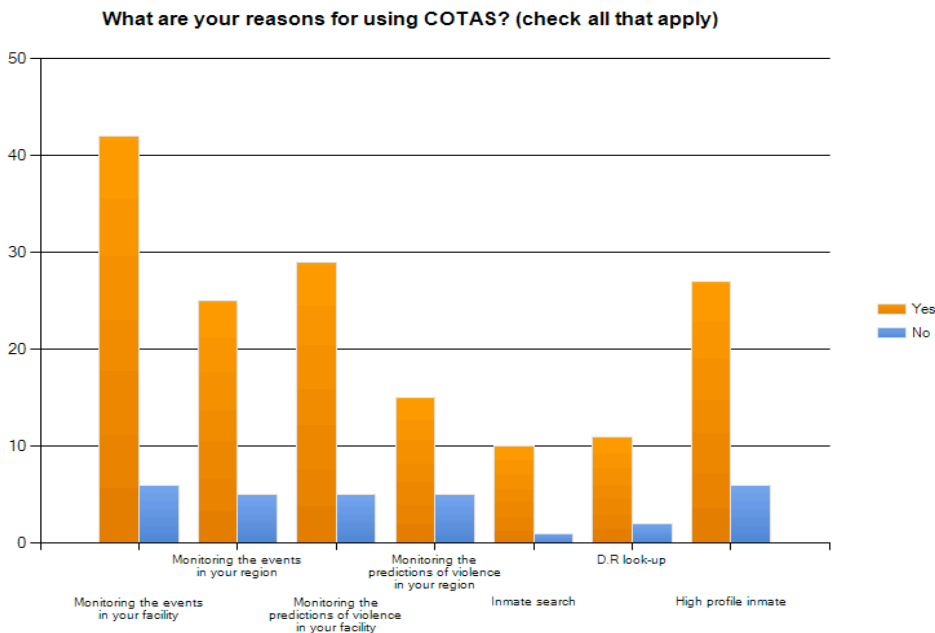


Figure 3: Use of Region Gauges by Respondents who Understand the Color Distinctions

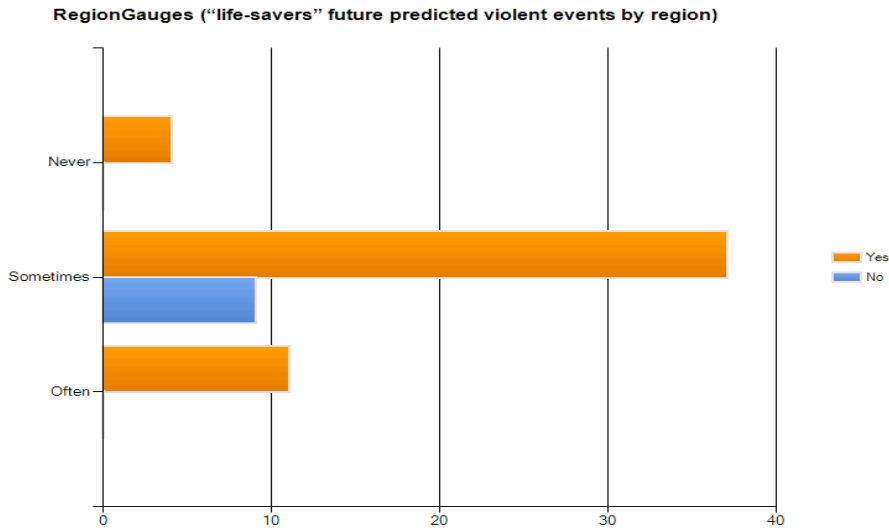
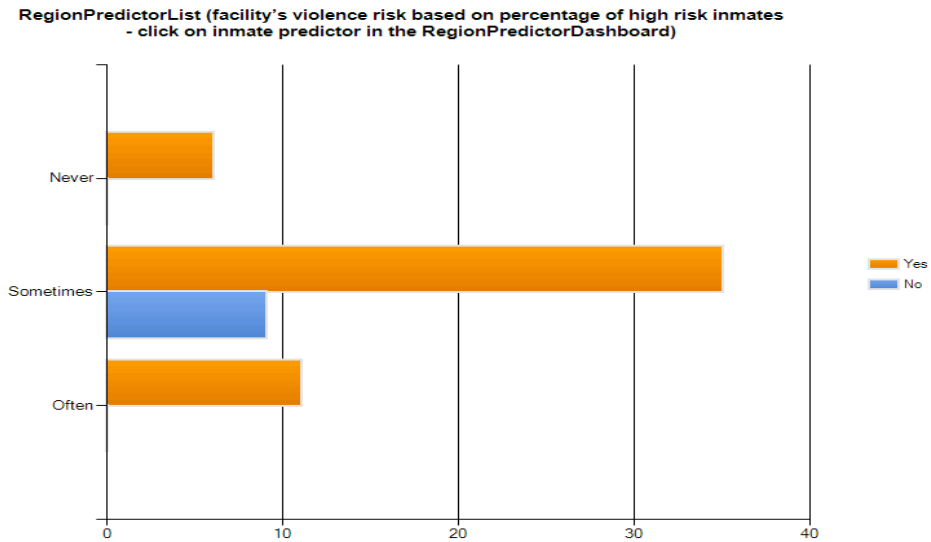


Figure 4: Use of the Inmate Predictor by Respondents who Understand the Color Distinctions



As can be seen in Figures 2, 3, and 4, users who understand the distinctions are more likely to make full use of all of the functions in COTAS. Furthermore, those who understand the distinctions are also more likely to use the prediction screens, especially the inmate predictor screen, where understanding the distinctions is paramount.

This outcome of this question also affects how useful the respondents found the Prediction Dashboard. As depicted in Figure 5, users who understand the distinctions find the predictive functions of COTAS very useful. In contrast, there is a higher concentration of those who do not understand the distinctions among users who do not find the Prediction Dashboard useful. This is also demonstrated in Figure 6 which

indicates the level of reported usefulness of the Inmate Predictor Function of the Predictive Screen. There is a similar pattern when examining the portions of the Prediction Dashboard that are most used by respondents. Finally, this question affected the overall goal of COTAS, which is to aid in the ability to predict the occurrence of violent events. As Figure 7 demonstrates, all respondents who “strongly agreed” to that statement understood the distinctions. Furthermore, there is a much larger presence of those who did not understand the distinctions among those who either “agreed” or “did not agree.”

Figure 5: Usefulness of the Prediction Dashboard by the Level of Understanding of the Color Code Distinctions

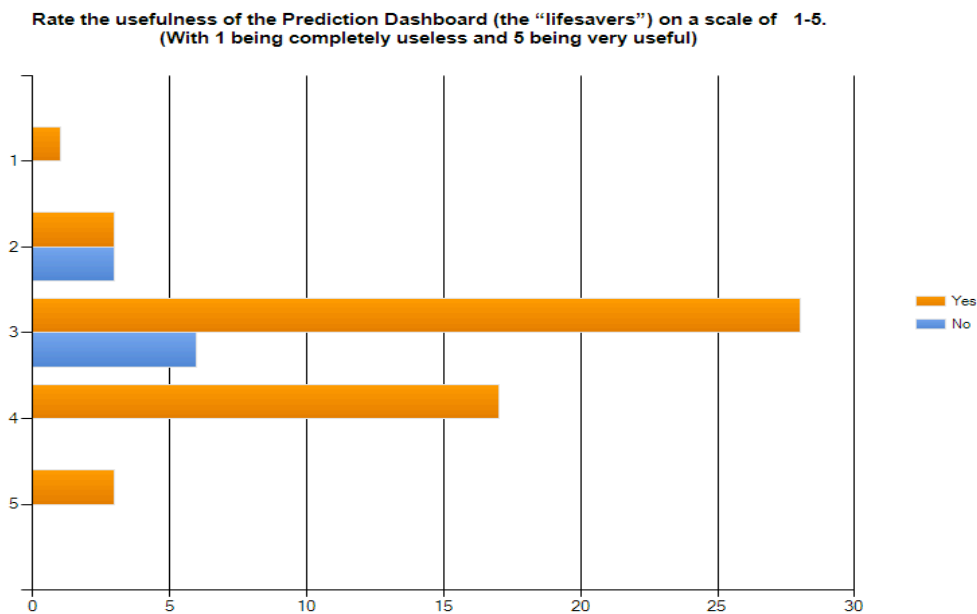


Figure 6: Usefulness of the Inmate Predictor Function (on the Prediction Dashboard) by the Level of Understanding of the Color Code Distinctions

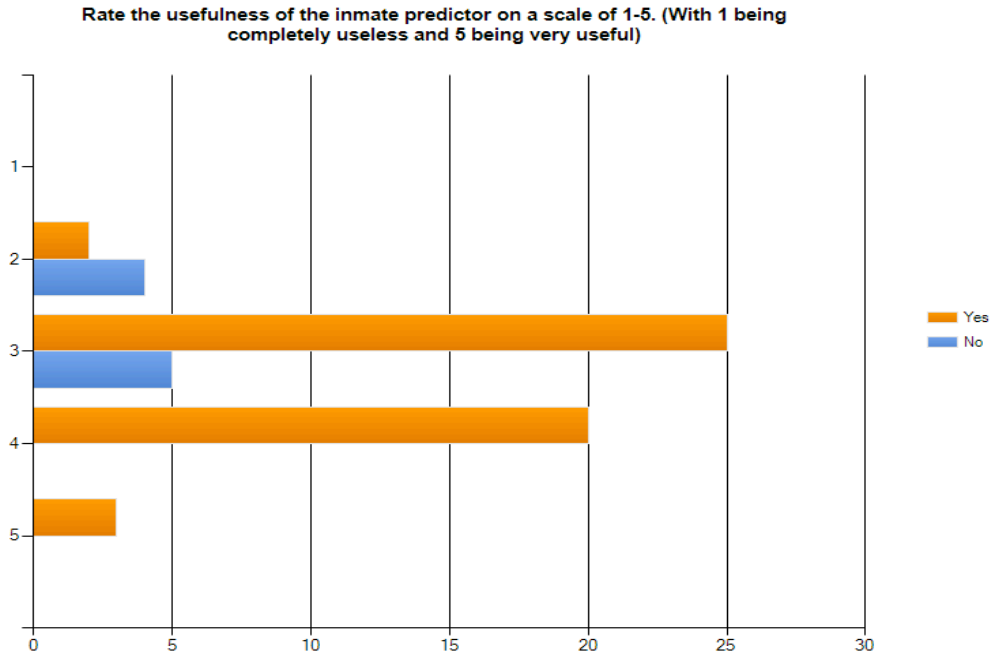
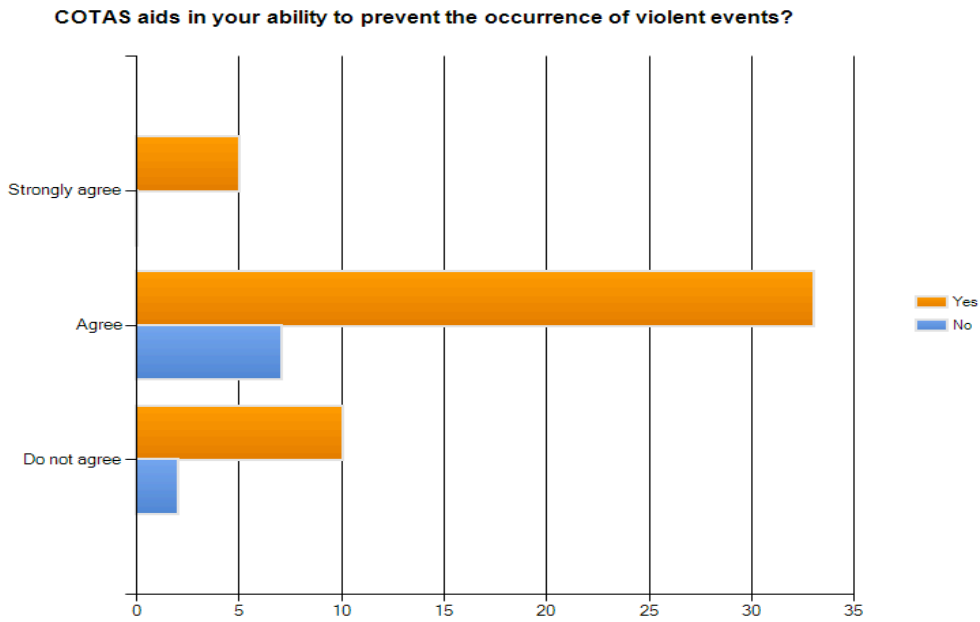


Figure 7: COTAS Aids in Ability to Prevent Violent Events by Users Who Understand the Color Code Distinctions



Were you aware that facilities are ranked depending on category (work release centers, work camps, and jails, female institutions, etc.)?

Responses to this question were informative regarding the usage, frequency, and comprehension of the various functions and screens of COTAS. As indicated by the pilot study, there was some confusion surrounding the categorization process in the RegionDashboard. This question was included further explore this issue and determine whether or not the respondents understood the categorization process. Project staff were able to identify the impact of the responses to this question on other questions.

As Figure 8 illustrates, there is no pattern among specific job titles/positions with regard to awareness of the ranking of the facilities' using color-coded categories (different types of facilities have different thresholds that determine whether they are categorized into red, yellow, or green categories). Understanding the coding system is important when assessing the usage of the RegionDashboard which ranks each facility within a specific region into red, yellow, and green categories. Figure 9 illustrates that the understanding of the ranking system is very important when it comes to using this screen. As Figure 9 depicts, the highest concentration of those who do not understand the ranking system indicated that they "never" use the RegionDashboard. Figures 10 and 11 summarize the responses regarding the accuracy and usefulness of the RegionDashboard. Figure 10, which displays the perceived accuracy of the RegionDashboard, indicates a somewhat random pattern. The high response rate of "do not know" may be an indication of a lack of understanding specifically what the survey question was asking. Figure 11 indicated a similar lack of a definitive pattern or trend. Perhaps, users' lack of an understanding of the facility ranking system is less important than users' lack of an understanding of the color coding distinctions for the inmate predictor (discussed with the previous question).

Figure 8: Job Title/Position by Awareness of the Facility Ranking System

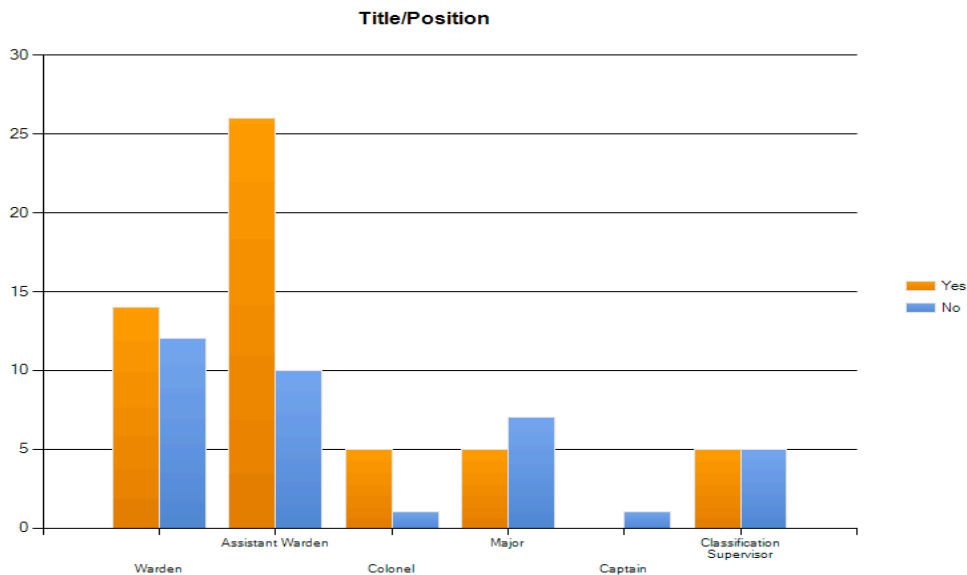


Figure 9: Use of the RegionDashboard by Awareness of the Facility Ranking System

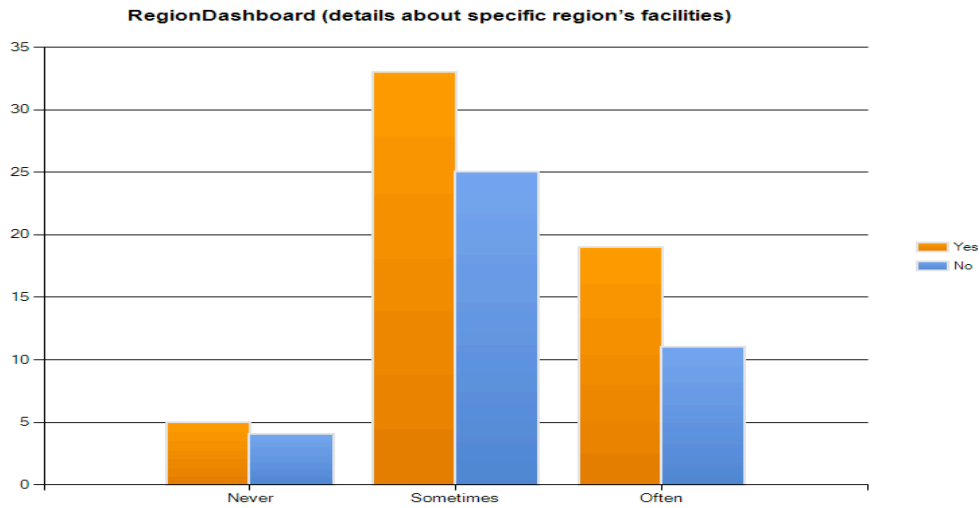


Figure 10: Accuracy of the RegionDashboard by Awareness of the Facility Ranking System

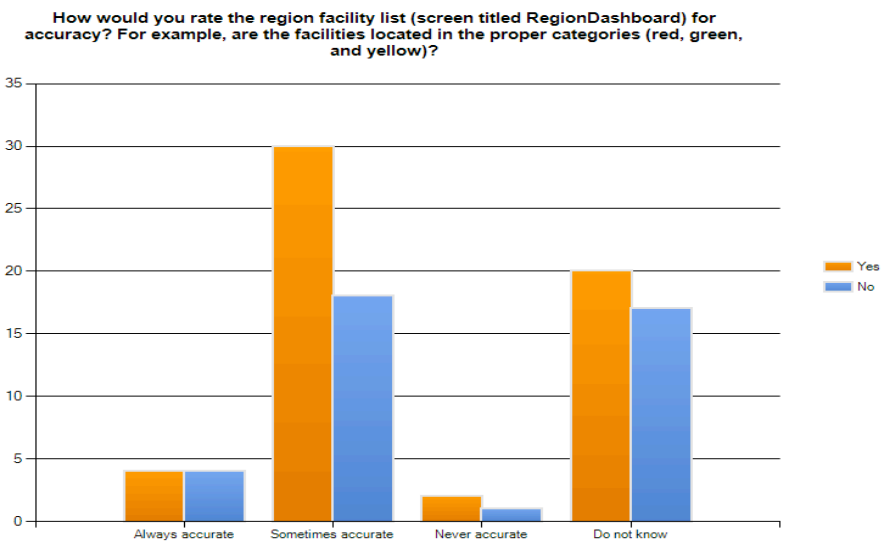
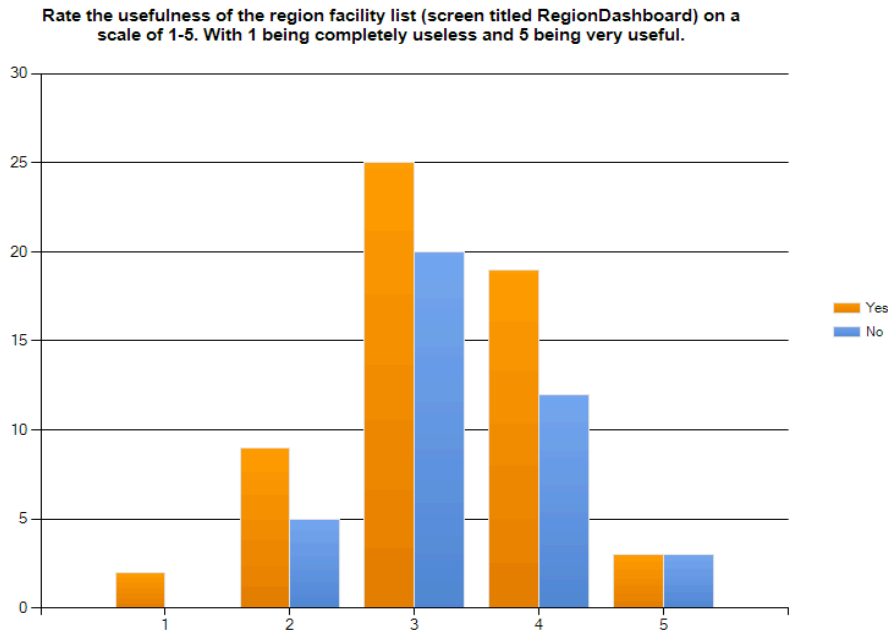


Figure 11: Usefulness of RegionDashboard by Awareness of the Facility Ranking System



Limitations

There were two limitations to this user survey. First, while the response rate was high (55%) considering that the respondents had only one week to respond to the survey, a higher response rate may improve the generalizability of the results. However, due to the time constraints, allowing additional time for respondents was not feasible. Second, the survey was dispersed by the Assistant Secretary of Institutions Russell Hosford. It would have been preferable to disseminate the survey notice and link from the independent evaluator/validator, the FSU research team. Having the survey email come from the Assistant Secretary of Institutions may have introduced some level of institutional pressure or loyalty by respondents. However, survey responses did not present any evidence of bias. DOC suggested using this strategy and, given the limited time to respond, researchers agreed in the hopes that it would increase the response rate. The population of professionals targeted by the survey is an extremely busy sector of Florida's correctional institutions and an email from an unknown or unrecognizable source (e.g. surveymonkey.com or FSU) requesting their participation in a 20 minute survey may not yielded such a response rate without encouragement from headquarters or a supervisor. However, to avoid potential biases associated with the survey email coming from DOC (their supervisor) the FSU researchers drafted the content of Assistant Secretary of Institutions Russell Hosford's email.

Summary and Discussion

The following concluding statements summarize the salient findings derived from the survey administered to DOC personnel relating to their use of and experience with COTAS.

- **Staff usage of COTAS:** A significant portion (42%) of DOC personnel who have access to COTAS do *not* use this tool. The top three reasons cited for not using COTAS are: not knowing they had access, not enough training to navigate COTAS, and COTAS is not viewed as useful because the same information is obtained elsewhere.
- **Frequency of COTAS usage:** Staff who reported that they use COTAS have approximately 21 months of experience with it and use it on a relatively frequent basis (once or twice per week).
- **COTAS screens/functions used:** The various COTAS functions are used relatively evenly.
- **The most popular COTAS functions:** to monitor the events in their facility and monitor events in their region.
- **The least popular COTAS functions:** disciplinary referrals (D.R.) lookups and inmate searches.
- **Region Dashboard screen:** Reported as somewhat accurate and somewhat/moderately useful.
- **Facility Monitor screen:** Level of accuracy reported as unknown, yet, most users considered it somewhat/moderately useful.
- **Prediction Dashboard:** Most users indicated they use it “sometimes” (60%); more than one-third (38%) of respondents reported never using it, and the majority (56%) of nonusers indicated the reason to be because it was unhelpful. Users indicated it was fairly useful.
- **Usefulness of inmate- versus facility-level predictors of violence:** About one-half of COTAS users reported the predictions of violence at the inmate- and facility-level to be moderately useful, 49% and 54%, respectively.
- **Red, yellow, and green inmate predictors of violence categories:** Majority of users understood the distinctions between the categories (85%) and three-quarters of respondents believe the distinctions are accurate.
- **Does COTAS aid in preventing violence?:** Approximately seven in ten COTAS users agree or strongly agree that COTAS helps prevent violence in their facilities.
- **Would COTAS be Improved with an Updated User’s Manual?:** Ninety percent of COTAS users agreed with this statement.⁶

⁶ The second, updated user’s manual was available to FSU, but had yet to be disseminated to DOC administrators in the field. The updated user’s manual was more comprehensive than the original but did not address the confusion indicated by the respondents in the Filtered Responses section. The addition of an explanation of the distinctions and the ranking system would be beneficial to the updated user’s manual before dissemination to DOC staff and administration.

Chapter 5

Summary and Recommendations

COTAS brings together DOC's large collection of historical and real-time data on inmate violence and to make it available to facility administrators in a concise and user-friendly web-based dashboard. Facility administrators can use this information to make decisions that may result in decreased injury to staff and inmates, a greater ability to maintain order in the facilities, and cost-savings to the state. As discussed in Chapter 4, a majority (83.5%) of facility administrators who responded to the survey currently use COTAS. This chapter summarizes the validation findings and presents recommendations to improve the accuracy and utility of COTAS. *Chapter 6 of this report presents the methodology, findings, and recommendations from the COTAS Software Validation that was conducted by ELENCO, Inc.*

Summary

Chapters 1 through 5 of this report describe COTAS, summarize the validation methods and findings, and present findings from the user survey. The validation began in February 2011 and was completed in June 2011. The COTAS validation examined the predictive measures for violence; the programming and software implementation; and the ability of COTAS to provide timely, accurate, and useful information to correctional administrators.

Overall results indicate that COTAS provides correctional administrators with useful information regarding the prevalence of violent and non-violent events within DOC facilities. The models predicting inmates' involvement in violent events, the documentation of the software ETL procedures, and the tests of data accuracy are areas that warrant dedicated attention to improve the accuracy and functionality of the system. Another key concern is the need for a data dictionary or codebook for all of the data used by COTAS. The remainder of chapter presents recommendations to improve the accuracy and utility of COTAS.

Recommendations

Recommendations are presented in the following order:

- Improvements regarding the user experience as reported through the COTAS user survey
- Improvements to the user interface
- Improvements regarding the predictive modeling and accuracy

Note: Chapter 6, COTAS Software Validation Report, is the self-contained report on the validation of the COTAS software which includes recommendations for software improvements.

Recommendations: The User Experience

While the results from this survey indicate that COTAS is being used throughout the state with some regularity, these results also highlight a number of issues with the use and understanding of COTAS by correctional officials and personnel. *Overall the greatest area of concern for COTAS users is the confusion and misunderstanding about exactly what COTAS does (purpose), what it is capable of doing (functionality), and who has access to it (availability).* The following recommendations reflect issues revealed through the survey administration.

Recommendation #1: Many of the problems associated with the user's experience can be addressed by an updated, expanded user's manual that is comprehensively disseminating it to correctional institutions throughout the state. The manual needs to provide complete explanations for the establishment of the thresholds (red, yellow, and green), the rationale for the thresholds, and the rationale for their significance. Also, the manual should clearly explain the distinctions for ranking facilities and, in general, provide a more clear and comprehensive explanation for COTAS's functions. DOC currently has two user's manuals: an initial manual and a not yet disseminated revised manual. However, the revised manual should be updated based on findings from this validation and pilot tested in the field.

Recommendation #2: DOC should make appropriate training opportunities available. Many respondents reported they had never been properly trained and indicated that they would like to be trained to facilitate their ability to navigate COTAS. Perhaps voluntary training sessions could be offered at each parent facility along with the dissemination of a new, updated user's manual.

Recommendation #3: Prior to updating the manual and developing training, DOC should coalesce on the primary purpose of COTAS, the desired uses/functions, and designate staff to provide "help desk" type assistance for users. Also, DOC should monitor the usage/traffic, track error reports and inaccurate data reports, and periodically solicit feedback from COTAS users in the field.

Recommendations: The User Interface

As the research team began the validation process, they documented irregularities with the interface.

Recommendation #4: The first issue involved the various links contained on the screens in COTAS. As is common with many websites, there were occurrences of selecting links and being directed to a website that was not functioning correctly. The nonfunctioning or broken links were not systematic; they were randomly dispersed throughout the COTAS screens.

Recommendation #5: The second issue involved the columns on the side of the RegionDashboard screen. Occasionally the numbers in the columns did not add up accurately, indicating to project staff that certain events were either not being counted or were being counted more than once. For example, an escape was listed as an escape and a non-violent disciplinary referral.

Recommendation #6: The third issue involved the inclusion of institutions in COTAS that are no longer operational. There were several institutions that had been shut down yet they remained on the list in the COTAS interface. This may lead to the miscalculation of numerous variables. For example, it appears that each inmate is a member of their own gang when in reality it is zero inmates in zero gangs which yields a percentage of 100 for gang members in that facility. As a result, these facilities appear to be the most dangerous when in reality they are actually no longer in existence.

Recommendation #7: DOC should thoroughly review the interface, re-pilot test it, and revise the code to correct these abnormalities. Further, it would be beneficial if DOC routinely or periodically sought input and feedback from users regarding improvements to the interface (e.g., additional screens, different layout, different presentation, options for customizing a screen, creating/saving reports).

Recommendations #8, 9, 10, and 11 address the use of color-coding of events and facilities in COTAS. There appears to be substantial confusion over the thresholds and the meanings of the color distinctions. While the color-coding of inmates, events, and facilities over the last 30 days and risk scores of inmates as “green,” “yellow,” and “red” was most likely designed to be intuitive to users (e.g., low, medium, high); however, without explanations about meanings, rationales for the thresholds, and consistency across all uses of color-coding—it probably is not very intuitive for users. However, color-coding is used throughout the user interface and the colors do not have intuitively-consistent meanings in the various contexts. It may be difficult for users to follow the distinct meaning of each color code. For example, the color-coding for events that have occurred in the last 30 days at the facility level is based on thresholds, which are weighted differently based on the facility type and category. Thus, each assessment is unique to a given facility. It would be helpful for users to know the specific threshold points that apply to their facility. Additionally, color-coding for inmates is displayed in aggregate in the “lifesaver” on the bottom of the RegionGauges screen and as bars on the RegionPredictorList Screen; however, only the risk scores are displayed on the PredictorSummaryByDorm screen. It would be helpful for users to have access to the inmate information (e.g., name, DC number, or risk score) by color category as currently available at the region and facility levels.

Recommendation #8: Remove the color-coding on the number of events at the regional level presented on the gauges on the top of the RegionGauges screen. The color coding that ranks the number of facilities with at least one violent event

displayed as red (top of the RegionGauges screen), does not provide any interpretable information. Indeed, given that regions vary in terms of the number of facilities, it may not make sense to apply a uniform metric based on a count of facilities. Providing facilities administrators with detailed statistics regarding events (both violent and non-violent) that have occurred within their facility is one of the major strengths of COTAS. According to the COTAS users' survey, most administrators found this information to be quite helpful. However, it would be beneficial to provide the user with a detailed explanation of the color-coding systems and threshold both in the users guide and on the COTAS interface. Additionally, threshold should be updated on a regular bases to reflect current trends in the entire DOC system.

Recommendation #9: Provide users with explanations of the color-coding systems and thresholds for inmates' involvement in events over the past 30 days at the facility level.

Recommendation #10: Update color-coding thresholds for inmates' involvement in events over the past 30 days at the facility level to reflect current data trends.

Recommendation #11: Provide color-coding of individual inmates on the PedictorSummaryByDorm screen. Inmates are assigned an inmate risk score from the COTAS algorithms and the scores are aggregated together into color-coded-categories at the facility and regional levels. It may be helpful to users to view these color-categories when looking at a list of inmates or an inmate profile.

Recommendation #12: Present data trends as changes in the rate of inmates involved in events, not the number of inmates involved in events, on the RegionGraph Screen and the EventGraph Screens. Changes in facilities' population size may also increase or decrease the number of inmates involved in violent events. Thus, reports of changes in trends over time should be based on rates not counts. For example, a rate might be the number of inmates involved in violent events per 1,000 inmates.

Predictive Modeling Recommendations

The recommendations included in this section were derived from the assessment of how the predictive models were generated for COTAS, how these models are used to make predictions of inmate violence, and how these results are presented to COTAS users.

Recommendation #13: BRDA should be more actively involved in the creation and review of statistical modeling techniques employed by future versions of COTAS and, therefore, more involved with overseeing the work of the subcontractor, *Idea*.

Recommendation #14: Create a codebook or data dictionary of variables in the model and data warehouse, and provide a clear description of the predictive model, including coefficients, measures of goodness of fit, and diagnostics. Additional problems with the COTAS algorithm may have been identified if

contractors from *Idea* had provided clear documentation regarding model selection, algorithm diagnostics, and a codebook or data dictionary.

Recommendation #15: Use the actual dates of events and have the Microsoft Logistic Regression Algorithm model the predicted probability of inmates' involvement in violent events within the following 30 days, and update the existing algorithm to reflect these changes. Many of the problems with the COTAS predictive models result from the fact that the dates of events are changed from the actual date to the first day of the month. As a result, the time-order distinction is lost distinguishing events (e.g., inability to discern time order for a bed change to solitary confinement and a violent event). Further, there is a loss of data when the last day of the month falls on a weekend.

Recommendation #16: Change the training data-set from static to dynamic with a "moving wall" of the previous 12 months of inmate data. COTAS should take full advantage of the Microsoft Logistic Regression Algorithm by changing the training data-set from static to dynamic and allowing the Algorithm to adjust to long-term changes in data or data trends, which may affect the impact of predictor variables on the outcomes.

Chapter 6

Validation of the COTAS Software

Summary

The objective of the Software Validation effort is to evaluate the software implementation of the predictive measures for violence and thresholds, with respect to:

- The implementation of the predictive measures and thresholds
- The aggregation of the data into the data warehouse

Findings. The current COTAS implementation provides useful functionality to its user base. However, long-term sustainability and evolution will be impacted negatively by the current data warehouse design, lack of critical documentation and systematic testing, and by the techniques used for predictive modeling and data visualization. The current implementation does not fully meet its stated objectives.

Table 1. Summary Findings of COTAS Software Validation

| Score (0-5) | Area of Evaluation | Discussion |
|-------------|--|--|
| 3 | 1. Software Design: how well the software is designed and coded. | System design based on industry-standard tool but does not reflect best practices for data warehousing and business intelligence (forecasting). |
| 3 | 2. Software Implementation: how well the application is implemented. | Operational system has not been subjected to formal testing and end-user validation for accuracy of reports. |
| 4 | 3. User Documentation: accuracy, sufficiency and conformance to as-built system. | User documentation is adequate description of as-built system; explanation of dashboards is problematic. |
| 2 | 4. Software Transferability: extent to which the software can be transferred to and used by another agency in a cost effective manner. | Overall lack of documentation of system requirements (e.g., data dictionary) and design rationale will impede the effort to transfer system to another agency. |
| 3 | 5. Software Extensibility: extent to which new software features and functions can be added at a reasonable cost. | The system design is based on industry-standard tools which facilitate extensibility, but poor design and lack of documentation will make extending COTAS difficult. |

Recommendations. A prerequisite to expanding the COTAS system or its user base must include: (1) remediating deficiencies in the documentation of system requirements and design rationale, data warehouse design, predictive models, data visualization controls; (2) a re-design of the data warehouse dimensions and facts; (3) implementing data cleansing in the extract-transform-load (ETL) process from the staging area to the data warehouse; (4) performing formal testing of the ETL process and the reporting

system with industry-standard documentation (plans, procedures and results). Future work on COTAS should be done by contractors with extensive experience in data warehouse design and business intelligence.

Scope of Work

Validation of the software implementation of the predictive measures for violence and thresholds:

- The implementation of the predictive measures and thresholds
- The aggregation of the data into the data warehouse

Validation procedures for the software implementation and the aggregation of the data are divided into the five key tasks outlined below:

1. Software Design evaluation tasks: (e.g., how well is the software designed and coded?)

- Review of the system level design documentation
- Review of the application design and associated documentation
- Review of the application coding and associated documentation
- Review of the database design and associated documentation

2. Software Implementation evaluation tasks: (e.g., how well is the application implemented?)

- Review of the online response time
- Review of the software stability
- Review of the reported software errors and associated corrections
- Review of the timeliness of software availability

3. User Documentation evaluation tasks: (e.g., are the materials accurate, sufficient, and up to date?)

- Review of the user documentation
- Review of the user training materials

4. Software Transferability evaluation tasks: (e.g., can the software be transferred to and used by another agency in a cost effective manner?)

- Evaluate the overall design for use by other agencies
- Evaluate the software with respect to transferability to another agency

5. Software Extensibility evaluation tasks: (e.g., can new software features and functions be added for a reasonable cost?)

- Evaluate the software with respect to extensibility to include additional features and data sources.

The subcontractor will have access to the software code, documentation, and training materials related to the software design and data evaluation.

Validation Methodology

This section presents the methodology used to perform the COTAS software validation.

Dimensions of examination/validation:

- System Development Process
- Software Products
- Software Operation

Validation activities

- Document Inspection
- Operation
- Ad-hoc querying
- Exploratory testing

System Overview

The source for all data used by COTAS is the Offender Based Information System (OBIS) computer application currently installed at the Florida Department of Corrections (DOC) whose development began in 1981. The system is administered and maintained locally by DOC and contractor personnel.

The major areas of functional support within OBIS include:

- Inmate Custody Tracking
- Inmate Classification Tracking
- Inmate Banking with Interface to Canteen
- Inmate Housing Assignment Tracking
- Facility Population Tracking
- Transportation Scheduling
- Inmate Movement Tracking
- Release Date Computation
- Parole and Probation Supervision
- Court Ordered Payments
- Field Investigation Tracking
- Collection and Reporting of Health Services Statistics

The technical scope of OBIS is characterized by:

- 4.2 million lines of COBOL code
- Approximately 340 character based screens lacking graphical user interface features, even though they are viewed on PCs.
- Approximately 1150 “green bar” reports without any graphical support.
- Over 300 hierarchical Information Management System (IMS) database segments containing millions of rows of data that are incompatible with relational databases.

- Approximately 57 DB2 tables that have been created to ease the load on the IMS database.

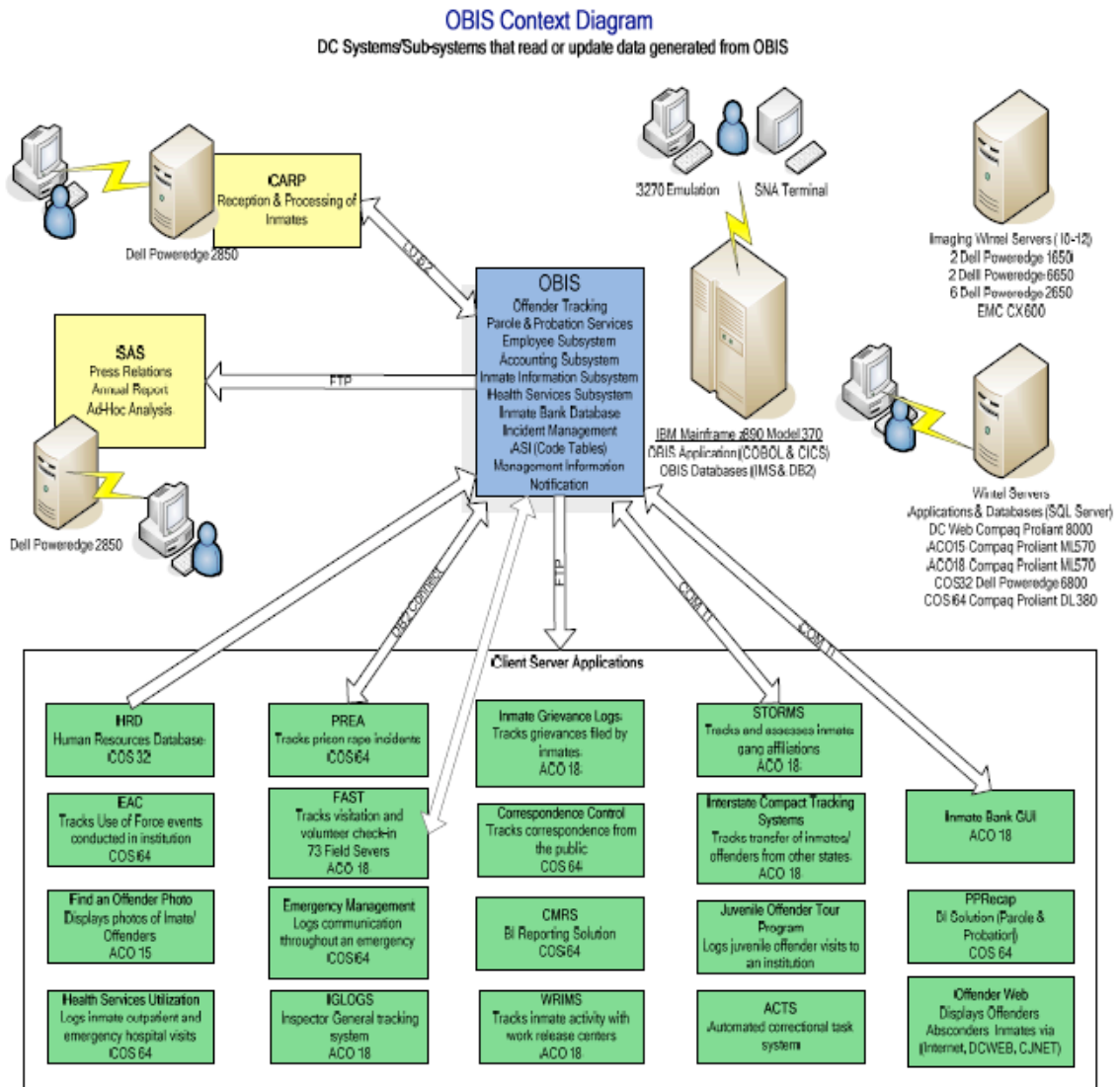


Figure 1. COTAS Data Source - OBIS

There are two separate extract-transform-load (ETL) processes that take the data from OBIS and move it into COTAS. The first ETL process moves the relevant data from OBIS into a staging database. The second ETL process then moves the data from the staging database into a separate dimensional database implementing the data warehouse. The dimensional database is queried directly for some of the COTAS reports. The remaining reports get their data from a cube that is built from the dimensional database. Additionally, the ETL process maintains audit logs of all transfer processes, and also populates a separate set of tables that contain information on any missing or erroneous data. An overview of the entire ETL process is given below.

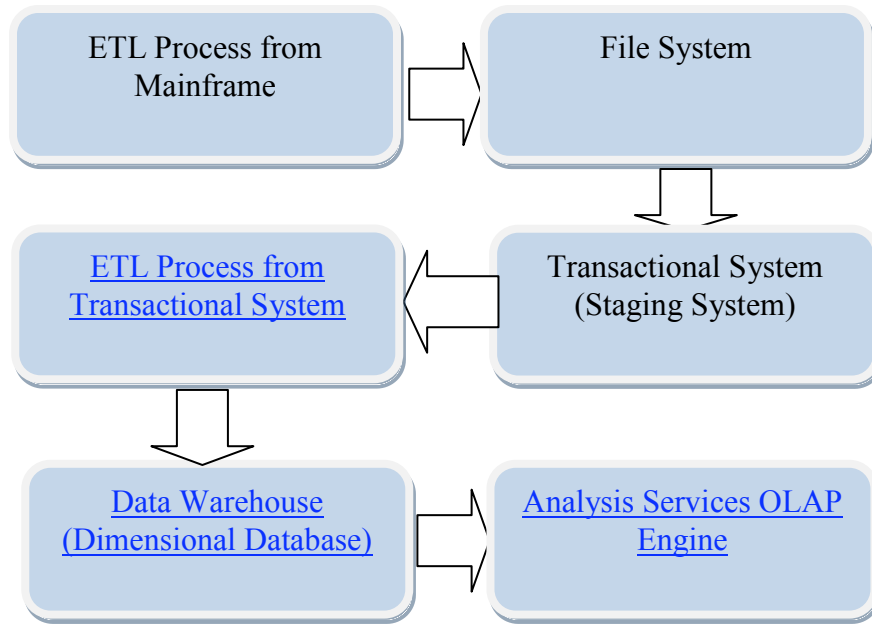


Figure 2. The COTAS Extract-Transform-Load (ETL) Process

The ETL processes are written as SQL Server Integration Services (SSIS) packages. The packages are stored in Microsoft Visual SourceSafe, a versioning system. The configuration information needed to execute the packages is stored externally, making it easy to transfer the packages from development to testing or production environments. The packages are executed nightly from Sunday through Thursday using SQL Server 2005.

Documentation is available for the entire ETL process. Each step in the process has its own package, and each package is described in the documentation. The documentation is given at a high level, but is fairly up-to-date and adequate to understand, troubleshoot and modify the existing ETL processes.

As implemented, the first stage of the ETL process (from OBIS to the staging database) is understandable, maintainable and generally conforms to accepted best practices. However, the ETL process from the staging area into the data warehouse omits a key practice of data warehousing, *data cleansing*. Data cleansing is performed to ensure the integrity of data warehouse contents, and will prevent duplicate data from being stored in the warehouse.

The Data Warehouse primarily consists of four components: (1) a dimensional database; (2) an Online Analytical Processing (OLAP) cube; (3) an analytical model that is used to predict the probability of inmate mischief; and (4) reports that comprise the user interface.

The dimensional database consists primarily of two types of tables: *fact* tables and *dimension* tables. The fact tables in a dimensional database contain things that are to be measured like number of items sold or total dollars spent. The COTAS dimensional

database contains two fact tables. The first fact table is used to count the number of negative events that involve inmates, such as fights with other inmates, breaking facility rules, escapes, etc. These events can be broadly classified into two types: violent and non-violent events. The second fact table measures the number of inmates and the number of gangs at a particular facility.

The dimension tables contain information that can be used to classify the facts. Restated, the dimension tables contain the information that is used to slice and dice the numbers represented by the fact tables. Examples of dimensions in COTAS include inmate information such as race, age and gender, type of event such as grievance, fight or escape, and facility information like type of facility or inmate population.

The dimensional database is stored on a SQL 2005 server. The tools and languages (SQL Server 2005 and SQL) used to manage the operational database are the same as for a dimensional database. The difference is in the design of the tables. A relational database (e.g., the operational database) is designed to reduce data redundancy and give balanced read and write performance. Dimensional databases, on the other hand, are designed to respond as quickly as possible to queries: redundant data actually provides great benefits in a properly designed dimensional database. Dimensional databases are generally designed into a set of *star schemas*. Each star schema consists of a single fact table surrounded by the various dimension tables that relate to the measures stored in the fact table.

The OLAP cube is built directly from a subset of the dimensional database. An OLAP cube is fundamentally a collection of pre-calculated sums. The core of an OLAP cube is the set of the measures which are retrieved from the fact tables in the dimensional database. The cube also contains information from several of the dimension tables. The numbers from the fact table are then summed up for each combination of primary keys for the dimensions that are part of the cube. If a user wants to know how many escapes occurred in the third quarter of 2010 in Leon County, the answer has already been calculated and stored in the cube. This makes retrieving the answer nearly instantaneous.

The final component of the data warehouse is reporting. In all, 22 reports have been created. These reports get their data either from the dimensional database, the OLAP cube, or a combination of the two. The reports are all created using Microsoft Business Intelligence Studio, and are hosted on a SQL 2005 Server. The reports can be accessed and distributed in a number of ways. The users gain access to the reports from a set of web pages that are hosted on an IIS server. Once the user is viewing a report, the user can export a copy of the report to a PDF, Word Document or Excel document.

System Development Plan Evaluation

The software system development lifecycle defines the process used to develop a system and the intermediate software artifacts produced in each phase of the project. There are many ways to manage a software system development project, but there are best practices and general guidelines for assessing the project management plan in terms of (1) the allocation of the total effort (labor hours) across the phases of the project; (2) the types of software system development products (artifacts) that should be produced, reviewed and controlled throughout the development process.

The areas of evaluation for the system development process are given in Table 2, along with an overall score for each area with a synopsis of the reason for the score.

Table 2. Software Process/Management Plan Evaluation Summary

| Score (1-5) | Area of Evaluation | Discussion |
|-------------|---|---|
| 5 | Does the project management plan comply with industry practices? | Sound, industry standard management plan. |
| 5 | Does the system development lifecycle define phases/tasks that comply with industry practices? | Industry standard development life cycle. |
| 4 | Does the planned effort allocated to each phase compare with industry practices? | Overall, adequate time was allotted for development phases. |
| 3 | Do the products from system development phases conform to industry practices? | Major products related to reviews and testing were not available (delivered by subcontractor) for evaluation. Significant gaps in the documentation. |
| 1 | Are those products saved in a controlled software configuration, as a permanent record of the project history? | The software configuration (collection of all system development products) is not under the control of the procurer. Implementation and design documents exist, but there is no associated revision history documenting review processes, debugging and system evolution. |
| 0 | Are industry standard processes used to verify that the system development products correctly reflect the system being built? | The set of reviews is reasonable; however, no records exist of the outcomes of reviews and testing. |

The system development process defined in the project plan conforms to industry practice. It is not clear that the process was followed during system development by the subcontractor, given the significant gaps in the set of software system development products, most notably the absence of requirements documentation and evidence of testing, and incomplete design documentation.

Effort Allocation

Based on the project plan in the “COTAS Timelines” document, the planned project duration was 16 months, with a total of 7600 labor hours (3.8 labor-years). Table 3(a) shows that 76% of the project effort was devoted to system design and development, with

14% devoted to data analysis and modeling. These effort allocations fall within industry norms.

Table 3(a). Planned Project Effort

| | Hours | Pct |
|---------------------------|-------|--------|
| Project management | 149 | 1.96% |
| Data analysis/modeling | 1075 | 14.14% |
| System design/development | 5792 | 76.21% |
| Operation | 584 | 7.68% |
| Total | 7600 | 100% |

Table 3(b) shows the allocation of effort for system design and development. The allocation of nearly 35% to requirements analysis is justified because COTAS is an unprecedented (first of its kind) system that involved a number of stakeholders. The 16% design effort allocation, though within acceptable ranges, may be low for a system that incorporates new technology (data warehousing, business intelligence and forecasting). System development and unit testing were allocated 35% of the total effort, again within industry norms. Finally, 10% for system testing is below industry norms.

Table 3(b). Planned System Development Effort

| | Hours | Pct |
|---------------------------------|-------|--------|
| Proposal | 32 | 0.55% |
| Requirements Analysis | 2024 | 34.94% |
| Planning | 64 | 1.10% |
| Design | 952 | 16.44% |
| Development(testing excluded) | 1272 | 21.96% |
| Unit Testing | 768 | |
| Inspections | 56 | |
| System Test Plan | 24 | |
| Acceptance Test Plan | 8 | |
| Total Development Testing | 856 | 14.78% |
| System Testing | 592 | 10.22% |
| TOTAL TESTING | 1448 | 25.00% |
| Total System Design/Development | 5792 | 100% |

Discussion

The reviewers could not determine the extent to which the project plan was followed, but based on informal discussions with DOC personnel, the reviewers intuit that a negative synergy may have occurred as a result of turnover of (DOC or contractor) project staff or the absence of a project champion. Given the fact that project was a first of its kind, positive synergy between the DOC and the contractor would have been critical to navigate successfully the many unknowns that surface during such a project. The fact that many of the industry-standard project deliverables could not be found suggests that much

of the project effort was spent exploring the problem and prototyping solutions. Focusing effort on these activities is common for first-of-a-kind systems.

Software Design Evaluation

The focus of the software design evaluation included the following:

1. Software Design evaluation tasks: (e.g., how well is the software designed and coded?)
 - 1a) Review system level design documentation
 - 1b) Review application design and associated documentation
 - 1c) Review application coding and associated documentation
 - 1d) Review database design and associated documentation

Table 4. Software Design Evaluation

| | |
|-----------------|--|
| <i>Approach</i> | <p>An extensive review of system design documentation and the system implementation focused on these criteria:</p> <ol style="list-style-type: none"> 1. Use of best practices in the data warehousing/mining design; 2. Use of industry standard data warehousing/mining tools; 3. Use of best practices of visual user interface design (visualization); 4. Completeness of the documentation; 5. Efficacy of the data visualization (dashboards, pie charts) in the user interface; 6. Appropriateness/correctness of the predictive models. |
| <i>Findings</i> | <ol style="list-style-type: none"> 1. The 2-stage data warehouse build (ETL) process, as implemented, is understandable, but does not perform <i>data cleansing</i> second stage, a departure from industry-standard practices. 2. No single formal list of system requirements was found; separate COTAS documents imply requirements. 3. Inadequate documentation to support system maintenance: complete annotated list of the stored procedures; data design (staging database, warehouse) using industry standard notations such as entity-relationship diagrams (ERDs), data dictionaries; warehouse dimensions and facts. 4. Non-standard design of the data warehouse dimensions and facts. 5. Data visualization wastes screen space; not clear what information the graphics are intended to convey visually; makes it hard to drill down. 6. Inconsistent user interface design for drill-downs. For example, the Non-Violent Disciplinary report contains collapsible sections. The main issue is that each section contains exactly one entry. When the report is first displayed, the vast majority of the display is simply empty white space. To see the details of an entry the user must click the small plus next to that line, which has been done for the third entry on the report above. To see the details for the entire report, this must be done <u>for each line</u>. 7. A careful usability study should be conducted to ensure that users can understand/comprehend the visual output and that users can work efficiently (otherwise, they may avoid using tiresome features). 8. The entire user interface needs to be revisited. The goals should be to make the information as accessible and understandable as possible in the shortest possible period of time. |

There are several areas in which the available documentation is inadequate. First, no documentation exists for the staging database that holds the data extracted from the source systems. A relational diagram that shows the tables, the columns and their types and the relationships that exist between the tables should be created. Second, and more critical, is that no data dictionary exists that explains the purpose of each table and each column in business terms. Third, a great deal of the logic that is used to transform the data for loading into the dimensional databases is contained in stored procedures stored in the staging database. Documentation of stored procedures is lacking: at minimum, a list of the stored procedures along with a synopsis of what each does should also be provided.

For each dimension, the diagram should include a clear description of how changes to the information stored in the dimension, known as slowly changing dimensions, are handled. If a facility name is changed at some point, for example, how will this be handled in the dimensional database? There are two major approaches to handling this type of change. The first, a type 1, is to simply replace all occurrences of the old name with the new name throughout the dimensional database. The second, a type 2, is to store both versions of the name. Each dimension (column in the dimension table) should have either a 1 or a 2 after to indicate how changes are handled.

Software Implementation Evaluation

The focus of the software implementation evaluation included the following:

- 2. Software Implementation evaluation tasks: (e.g., how well is the application implemented?)
 - 2a) Review online response time
 - 2b) Review software stability
 - 2c) Review reported software errors and associated corrections
 - 2d) Review timeliness of software availability

Table 5. Software Implementation Evaluation

| | |
|-----------------|--|
| <i>Approach</i> | Work on this task requires access to the COTAS development environment for first-hand observation of system behavior and response. An informal exploration of the system quickly led to several instances of invalid stored data (integrity) and ambiguous/incorrect query responses: the root cause of these inconsistencies may lie in the data warehouse design or the software (stored procedures) used to load the warehouse. Historical data on software stability, revision history and system availability was not available for inspection: anecdotal evidence suggests that such data do not exist. |
| <i>Findings</i> | <ol style="list-style-type: none"> 1. Data anomalies were found in which the certain events were misclassified (e.g., a DR on an escape is treated as a violent event), or miscounted (a DR on an escape is also counted as an escape), etc. 2. Certain invalid dates have been found in the data warehouse. 3. Because these anomalies were discovered by informal exploration of the system, one might assume that more extensive formal testing would find a number of serious errors that could compromise the validity of decisions made using erroneous data. 4. Caveat: relying on end users to discover errors is not a substitute for formal, repeatable testing. End users tend to ignore errors that appear innocuous when, in fact, these small errors may compromise the validity of data in detailed and |

| | |
|--|--|
| | summary reports, and in predictive measures. |
|--|--|

The second concern is the lack of evidence that testing was performed. The system architecture only has two environments: a development environment and a production environment; there is no testing environment, which is unusual for a system of this size. No documentation of any testing was available; there were no testing plans and no testing results for any component of the system, including the ETL process.

The lack of evidence of testing is possibly the most significant finding, in that many of the anomalies discovered during the validation process would certainly have been caught prior to system release had testing been performed. There is no empirical evidence that the data stored in the data warehouse or displayed in COTAS reports accurately reflects what was stored in the operational system. There is not even evidence that the data was moved to the staging database correctly. Indeed, a quick survey of the data found several data anomalies.

The entire COTAS workflow process contains many data processing steps, as shown in Figure 1; errors in one step may propagate into subsequent steps. A series of well defined tests should have been defined, stored, and used to verify that each step is performed correctly under circumstances both favorable and unfavorable. Although one may infer that some informal process must have been performed, industry-standard test plans, test data and test results appear not to have been produced and retained.

Additionally, detailed tests of the stored procedures that perform database and warehouse extractions, updates and reporting should have been produced.

User Documentation Evaluation

The focus of the user documentation evaluation included the following:

- 3. User Documentation evaluation tasks: (e.g., are the materials accurate, sufficient and up-to-date?)
 - 3a) Review user documentation
 - 3b) Review user training materials

Table 6. User Documentation Evaluation

| | |
|-----------------|--|
| <i>Approach</i> | User and training documentation— <i>COTAS User Manual</i> and <i>COTAS Tutorial</i> —were reviewed for content (just reading the document) and for consistency with the implemented system (following the document while using the system). Disclaimer: no exhaustive “test” of the documentation was attempted. |
| <i>Findings</i> | The overall quality of the user documentation is good and consistent with the actual system behavior. The training/tutorial documentation was deficient in several areas: <ol style="list-style-type: none"> 1. The drill-down functionality appears to work as described in the documentation. 2. The explanation of the Region Gauges in the tutorial is adequate (details about the factors and thresholds that place the gauge in the red, green or yellow sections are omitted). 3. The explanation for the lifesaver donut-charts in the RegionGauges screen is not |

| | |
|--|--|
| | <p>clear.</p> <p>Recommendation: The business rules that determine the visual display (colors, “odometer” reading, and correct interpretation) are <u>system requirements</u> that should be documented more effectively, to ensure that the displays are used/interpreted correctly by end users, and to provide the basis for testing/validating correctness of system outputs.</p> |
|--|--|

Software Transferability Evaluation

The focus of the software transferability evaluation included the following:

4. Can the software be transferred to and used by another agency in a cost effective manner?
 - 4a) Evaluate overall design for use by other agencies
 - 4b) Evaluate software with respect to transferability to another agency

Table 7. Software Transferability Evaluation

| | |
|-----------------|--|
| <i>Approach</i> | <p>Expanding the COTAS agency user base will involve modifying the current design and implementation to accommodate new users and requirements specific to each new agency. New requirements related to reporting and prediction must be compatible with existing requirements. Finally, an agency-specific extract-transform-load (ETL) process must be developed and tested for each new agency: under ideal circumstances, the ETL process would constitute the bulk of the transfer effort.</p> <p>Documentation for the existing COTAS system must be the basis for adapting the system to new agencies.</p> |
| <i>Findings</i> | <ol style="list-style-type: none"> 1. Deficiencies in the current documentation set should be remedied. 2. One area of documentation must be a concise definition of the predictive models and thresholds used to forecast violent incidents. 3. Formal acceptance testing (test plan, test requirements, test data sets) should be performed to establish a correctness baseline that can be used to monitor system evolution (regression testing). 4. Deficiencies in data warehouse design should be addressed; an improved design is likely to be more maintainable. 5. A long term approach that can decrease the effort/cost of adoption by a new agency is to implement a service-oriented architecture for the ETL process. |

Software Extensibility Evaluation

The focus of the software extensibility evaluation included the following:

5. Can new software features and functions be added for a reasonable cost?

5a) Evaluate the software with respect to extensibility to include additional features and data sources

Table 8. Software Extensibility Evaluation

| | |
|-----------------|--|
| <i>Approach</i> | The data warehouse documentation and design are the key work products used to extend COTAS features and data sources. Changes to data sources will require revisions to the data warehouse design and the warehouse loading (ETL) process. |
| <i>Findings</i> | <ol style="list-style-type: none"> 1. Documentation of data warehouse and ETL processes needs to be created before extensions are considered. 2. Before adding new features or data sources, the data warehouse needs to be redesigned to conform to best practices for dimensional modeling. 3. A formal system test should be developed before extending COTAS. This test can be used later for regression testing (i.e., testing to ensure that future changes do not destroy pre-existing functionality). 4. Given the current implementation of COTAS, maintenance or extension of the system by DOC personnel would be risky: significant training would be required in data warehousing, the use of the Microsoft tools, and predictive modeling. |

Recommendations

What Should Be Done to Conform to Industry Practices

1. Document the results from a complete requirements analysis (to serve as the basis for design, implementation and testing)
 - Build a data dictionary to establish terminology and usage rules for all data fields, tables and reports. (Note: **The confusion over “event” versus “report” would have been prevented by a clear definition.**) [See the “Data/Reporting Anomalies” section in the Software Validation Report Appendix/Exhibits.]
 - Develop Entity-Relationship-Diagrams (ERDs) for source and staging databases, and for data warehouse design
 - Develop a master list of functional requirements for data warehousing, addressing things like: data sources, relevant fields, *data cleansing* (e.g., removal of duplicate data), rationale for establishing dimension tables (e.g., the questions the warehouse will be used to answer) and the facts (what will be measured, and at what granularity?).
 - Develop a master list of functional requirements for reporting: what each report should contain, and should NOT contain; the types of errors that should be prevented.

2. Comprehensive design documentation to support future system maintenance, extension and transfer to other agencies (to serve as the basis for implementation and testing).
 - Document data warehouse dimension and fact models, to include: ERDs, fact granularity; rationale for choice of slowly-changing dimensions;
 - Expand ETL process documentation beyond merely printing the diagrams produced by Microsoft SSIS, to include decisions that may impact future changes (e.g., data sources, dimensions or facts);
 - Document stored procedures to include the function performed, error checking/handling performed, and rationale for choosing from alternative features/solutions (especially when a non-standard approach is taken). This documentation becomes the basis for testing.
3. Systematic testing process for the major system elements
 - Test documentation should include:
 - Test plan (specific objectives – tied to a requirement or component (e.g., stored procedure, ETL step, report);
 - Test data (data needed to set up a test; data describing expected results);
 - Test procedure – repeatable process for setting up the test, executing the test, and capturing the results from the test.
 - Unit Testing – test objectives should be based on the function performed and the error checking/handling required by:
 - Stored procedures in the ETL and report generation processes;
 - Step(s) in the ETL process;
 - COTAS reporting system
 - Report generation.
 - Testing the predictive measures/reports
 - System Testing – test objectives should be based on the end-to-end function performed by the overall system.
 - Acceptance testing to demonstrate that the system satisfies requirements – typically a confirmation (positive) test focused on correct data and usage
 - System test with destructive (negative) tests focused on incorrect data and/or usage
 - The system test becomes the basis for regression testing of future revisions to the system (adding new features, data sources)
 - Usability Testing - a careful usability study should be conducted:

- to ensure that users are able to understand/ comprehend the visual output;
- to establish that users can use all COTAS features;
- to measure user perception that COTAS increases their productivity (i.e., features do not require unnecessary steps, keystrokes or mouse clicks) ; and
- To ascertain DOC staff use COTAS reports to inform their decision making.

Recommendations for COTAS Project Management

1. DOC should provide a project champion to ensure regular and substantial two-way communication between the developer (contractor) and the DOC. The DOC champion should be stable throughout the duration of the project.
 - a. Failure to have a stable DOC champion may lead to bad assumptions by the contractor personnel when requests for clarification or direction are not answered in a timely manner.
 - b. The DOC champion needs to have management support in order ensure access to key stakeholders at critical junctures in the project.
2. DOC should insist on firm deliverables, which should be placed under configuration management at the appropriate stages in the project (e.g., requirements should be controlled after initial sign-off).
3. DOC should subject the deliverables to a systematic review process; the review process should include the process being used by the contractor (e.g., via a third party review such as a SQA audit)
 - a. Whenever possible, involve the DOC IT personnel who may be responsible for operating and maintaining the operational system
4. DOC should ensure that the contractor selected to extend COTAS has substantial experience in data warehousing and business intelligence.
 - a. It is not enough to use the right tools (Microsoft SQL Server SSIS): the tools must be used in a way that conforms to industry best practices.
 - b. Experience is especially critical to developing a system that is expected to have a long operational lifespan and a broad user base.
5. DOC should require that contractor personnel assigned to the project be stable throughout the duration of the project.

Software Validation Report - Appendix/Exhibits

Data/Reporting Anomalies

Exploratory querying of the data warehouse uncovered a number of data anomalies. The largest category of these is double (or multiple) counting of events – sometimes in multiple categories. It was beyond the scope of this validation project to conduct an exhaustive search for all existing data anomalies. The ease with which several blatant anomalies were discovered, coupled with the lack of evidence of testing, strongly suggest that other, possibly more serious, anomalies exist.

One type of anomaly that was easily discovered concerns inmate escapes. Escapes are actually a more frequent occurrence from low security facilities. Escapes are often as simple as the inmate getting in a parked car with a friend and driving off. The anomaly concerns the way in which the escapes are counted. Each escaped is often counted and reported as two or more separate events.

The root cause for this anomaly has to do with the way inmate escapes are handled in the OBIS system. They are stored in an escape table (or mainframe equivalent). Once the inmate is caught one or more disciplinary reports are filed. These disciplinary reports are the source of the second event and the source of the double counting; this is because the disciplinary reports are not correlated with the escapes. To make matters worse, all disciplinary reports of type “Escape” are automatically flagged as violent, when the vast majority of escapes are, in fact, non-violent events. This means that escapes from low security facilities are often reported to the user once as a violent event, and one or more times as a non-violent escape.

The largest contributing factor to this class of anomaly is that a key business idea was not well defined, (i.e. no data dictionary). The specific business idea that is the root cause of many of the duplicate data anomalies is the failure to clearly define the idea of an “**event**”. The various dashboards in use display counts of events – grouped into a number of classifications, i.e. violent, non-violent, escapes, etc. The issue is that the numbers shown are wrong: the same event is often counted two or more times.

- Escapes
 - These are generally non-violent events. The most frequent scenario is that an inmate at a work release center simply does not return.
 - These are counted once as non-violent escapes in the user interface.
 - The same escapes are also counted one or more times under the Violent Disciplinary Reports category.
 - In the year from 5/24/2010 to 5/24/2011, there were a total of 166 recorded escapes.
 - For that same time period, there were 217 Violent Disciplinary Reports generated for those escapes.
 - Escapes were over reported by 30%.
 - None of the 166 escapes was, in fact, violent.

- Note that the number shown on the dials on the Facility Dashboard do not match the number shown in the odometer.
- Disciplinary Reports
 - The source system contains a large number of duplicate disciplinary reports (DR). In a one year span from 5/24/2010 to 5/24/2011, there was a total of 111,811 DR.
 - Of those 2,520 (or 2.25%) were duplicates.
 - In the worst single case, duplicates of the same DR were counted eight (8) times.
 - These duplicates are included in the counts of events shown to users.
 - Note that the number shown on the dials on the Facility Dashboard do not match the number shown in the odometer.
 - Fights
 - When inmates fight, a DR gets generated for each inmate involved in the fight.
 - When counting the total number of violent events – a single fight will always be counted at least twice.
 - A single fight could be counted up to 12 times* (as in a fight on 3/26/2011) – once for each inmate involved in the fight.
 - In the year from 5/24/2010 – 5/24/2011 there were 2,605 fights reported but 5,220 fights were reported to the users (i.e. there were 5,220 DR generated)
- Investigations
 - A quick search did not yield any anomalies.
 - Note that the number shown on the dials on the Facility Dashboard do not match the number shown in the odometer.
- Use of Force
 - Found evidence that multiple Use of Force reports (UOF) were created and counted for the same event (see 2011-102-0076 & 2011-102-0078, 2011-102-0077 & 2011-102-0079).
 - Found evidence that the number of inmates involved in a UOF was misreported. (see 2011-102-0078 – a single inmate involved but two inmates involved was reported)
 - The number shown on the Use of Force odometer does not equal the number of use of force detail lines shown on the detail report.
 - Note the number shown on the dial on the Facility Dashboard does not match the number shown in the odometer.

Conclusions

- In general, COTAS is counting reports and presenting the counts to the user as the number of events.
- No evidence was found showing that an attempt was made to consolidate reports of different types (i.e. Disciplinary Reports, Escapes, Investigations, Use of Force).
- No evidence was found showing that effort was made to identify and/or remove duplicate reports.

- Because of this, the numbers reported do not accurately represent the number of events that have occurred.
- For some types of events (fights, escapes) the numbers reported are at least double the actual number.

Dimensional Modeling Anomalies

The design of the dimensional model clearly does not conform to industry best practices. The primary fact table used in this model is the Fact_Event table. A small subset of the most obvious issues with the fact tables (the list is not exhaustive) follows:

- The name suggests that the table stores facts about events.
 - Correspondingly, the granularity of the fact table could be assumed to be that of a single event, i.e., each row in the fact table represents a single event (e.g., an escape or a fight).
 - In actuality, each row represents a single entry or report, from one of six sources:
 - Escapes
 - Use of Force
 - Investigations
 - Field Grievances
 - Office Grievances
 - Disciplinary Reports
 - If the same event, a fight for example, is reported in three of these sources, i.e., three rows will exist in the fact table.
 - This indicates that the fact table is actually measuring *reports* of events – and at the very least is very poorly named.
- The fact table contains null dimensional key fields – a major violation of dimensional modeling.
 - The fact table contains exactly eleven dimensional keys (foreign keys into dimension tables):
 - Facility
 - Violent (yes or no)
 - EventType
 - DATE
 - TIME
 - FieldGrievance
 - DisciplinaryReports
 - OfficeGrievance
 - Investigations
 - UseOfForce
 - Escapes

- Five of the last six dimensional keys contain nulls for every row in the fact table. Restated: only one of the last six dimensional keys contains a value for every row – the other five are null.
- Given the granularity of the fact table as it exists, it is missing an obvious dimensional key.
 - The fact table does not contain a dimensional key to the offender dimension.
 - To discover who the offender was requires looking into one of the six dimensions by following whichever dimensional key is not null for each row (inefficient).

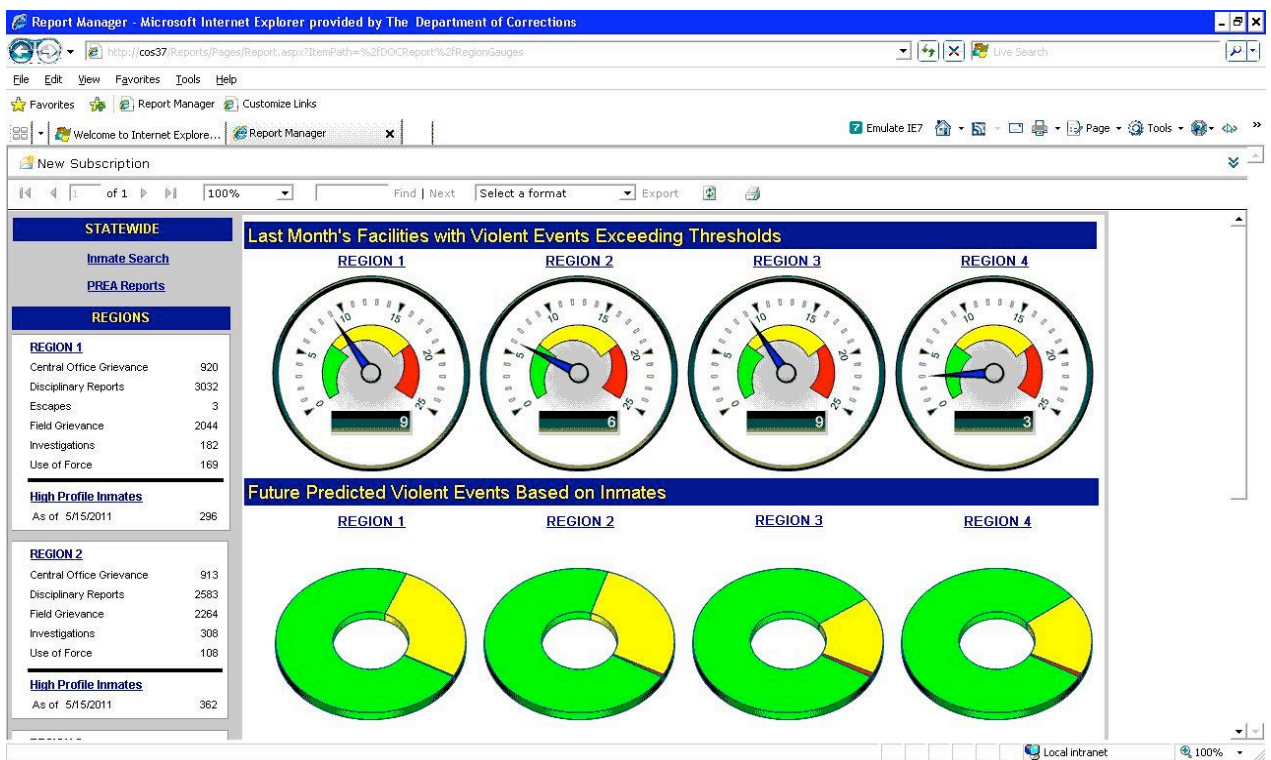
The dimension tables were not modeled properly. The dimensions contain copies of dimensional keys (which violates modeling conventions).

- dim_DisciplinaryReport
 - This dimension contains a copy of the Facility key
 - It also contains the NAME of the Facility – this information clearly belongs in the dim_Facility table ONLY.
 - The reason for this has to do with slowly changing dimensions. If the facility name were to change (which can happen) – all of the rows in the dim_DisciplinaryReport for that facility would need to be updated by either adding a second copy of each row – or by updating the facility name on each row.
- dim_Investigations
 - This also contains a copy of the facility id and facility name
- GrievanceFieldDetail (note the departure from the table naming convention)
 - This also contains a copy of the facility id and facility name
 - It also contains name of the inmate (or offender). An offender dimension (dim_Offender) exists. Storing the name of the offender in the GrievanceFieldDetail will cause the same slowly changing dimension issue as storing the facility name.
- GrievanceOfficeDetail
 - This also contains a copy of the facility id and facility name
 - It also contains the inmates (or offenders) name.
 - Unlike the earlier examples, in this dimension the inmate's name is stored as first and last name.
- dim_Escape
 - This contains a copy of the facility id
- UOFDetail (note the departure from the table naming convention)
 - This also contains a copy of the facility id and facility name

User Interface Concerns

The user interface for the COTAS system provides access to more than 20 reports. Some of the reports are laid out to resemble dashboards with a set of four gauges and doughnut charts, while most of the reports resemble traditional green bar reports. Navigation is accomplished by the user clicking on some report artifact, which then typically shows a related report with more detailed information.

The usability issues with the reports are numerous. Only a few of the most glaring issues are outlined in this section. A sample of the main dashboard retrieved on May 15, 2011 is shown below.



Fully 2/3 of the area of the dashboard has been wasted. Notice that fully 1/3 of the screen has been devoted to display four numbers: 9, 6, 9 and 3. The gauges used are very hard to read, which is why the actual numbers are also shown in the odometer of each gauge. This is space that could have been used to display additional information which currently does not fit on the screen.

Even worse are the four doughnut diagrams. These also occupy 1/3 of the screen. After studying these charts at length and reading the documentation supplied, the reviewers are still not certain what information these are meant to convey.

One problem with pie charts in general is that it is very difficult for people to accurately estimate the percentages represented by each section. The smaller the sections, the more

difficult it is for a user to accurately estimate the size of a slice. Doughnut charts make this task impossible by removing the middle, which contains the point where the lines intersect. Given these issues, it is clearly impossible to estimate the percentage of the green, yellow or red sections for any of the doughnut diagrams shown. In any case, pie charts and their derivatives are one of the most misused of charts; they are meant to allow users to compare parts of something to the whole. In this case, the heading for this section indicates that a probability will be shown: a pie chart is inappropriate.

For a detailed discussion of useful dashboard design and data visualization, the Steven Few book, *Information Dashboard Design: The Effective Visual Communication of Data* is highly recommended.

Another type of COTAS usability problem occurs in the detail report of non-violent disciplinary reports shown below.

The screenshot shows a web browser window titled 'Report Manager - Microsoft Internet Explorer provided by The Department of Corrections'. The address bar shows a URL starting with 'http://cos37/Reports/Pages/Report.aspx?ItemPath=%2FDOCReport%2FEventSummarybyDorm&ExecId=uo0heaajetuoyj45vz35Fhjn&PingId=fq5wlj'. The browser's address bar shows '#/FAMU:::37494765' and 'Report Manager'. The report content is as follows:

| Events for Facility: TALLAHASSEE W.R.C. | | | | | | | | Total Disciplinary Reports 16 | |
|---|--------|----------------------|------|-------------------------|---------------------|------------|---------------|-------------------------------|--|
| Events | Bed | Bed Mission | Dorm | Primary Work Assignment | Description | Log Number | Incident Date | Incident Time | |
| 168-110045 | | | | | | | 04/27/2011 | | |
| 168-110048 | | | | | | | 04/26/2011 | | |
| 168-110044 | H1109U | ADMINISTRATIVE CONF. | H | UNASSIGNED | DEVIATING VWO PERM. | 168-110044 | 04/25/2011 | 13.19.00 | |
| 168-110041 | | | | | | | 04/23/2011 | | |
| 168-110043 | | | | | | | 04/23/2011 | | |
| 168-110042 | | | | | | | 04/23/2011 | | |
| 168-110040 | | | | | | | 04/23/2011 | | |
| 168-110039 | | | | | | | 04/23/2011 | | |
| 168-110037 | | | | | | | 04/18/2011 | | |
| 168-110038 | | | | | | | 04/18/2011 | | |
| 168-110036 | | | | | | | 04/17/2011 | | |
| 168-110033 | | | | | | | 04/15/2011 | | |
| 168-110034 | | | | | | | 04/15/2011 | | |
| 168-110035 | | | | | | | 04/15/2011 | | |

This report contains collapsible sections. The main issue is that each section contains exactly one entry. When the report is first displayed, the majority of the display is simply empty white space. To see the details of an entry the user must click the small plus next to that line, which has been done for the third entry on the report above. To see the details for the entire report, this must be done for each line.

The entire user interface needs to be revisited. The goals should be to make the information as accessible and understandable as possible in the shortest possible period of time.

The goal of predicting violent incidents is ambitious: it is a hard problem to solve, and the solution is hard to explain and hard to visualize. It is important that this aspect of

COTAS use simplicity as a design constraint, as opposed to complex visuals whose meanings/interpretations are unclear to the user community.

Finally, predictive modeling is not a closed-form problem for which there is a single solution, but at a given point in time, a “best available” solution exists. Over time, other best-available models will emerge, informed by the empirical data in the data warehouse. COTAS must support the evolution of predictive models, including training for DOC personnel to understand, validate and upgrade models.