

The Annie E. Casey Foundation's Juvenile Detention Alternatives Initiative Data Analysis: Interim Progress Report

In March 2008, the Governor's Juvenile Justice Advisory Committee (GJJAC) contracted with the Washington State Center for Court Research (WSCCR), the research section of the Administrative Office of the Courts (AOC), to perform an evaluation, make recommendations, provide technical assistance, and conduct analyses of the Juvenile Detention Alternative Initiative (JDAI) in Washington State. Juvenile courts participating in JDAI are King, Pierce, Spokane, Whatcom, and Benton/Franklin.

This is an extensive study and this report is the third in a series, following the Data Capacity Assessment and Recommendation of Standards.¹ The Data Capacity report describes the sites' 1) JDAI data collection efforts, 2) analysis of JDAI data, and 3) production of JDAI-related reports; the Standards report details working group recommendations to standardize and reconcile site analysis and reporting practices. This Interim report describes the implementation plan, necessary activities, and site participation towards integrating the standards into analytical and reporting practices. Once adoption is complete, these recommendations will produce comparable measures and outcomes that support

¹ E. Valachovic, 2008. *The Annie E. Casey Foundation's Juvenile Detention Alternatives Initiative Data Analysis: Data Capacity Assessment*. Olympia, WA: Washington State AOC. and

E. Valachovic, 2008. *The Annie E. Casey Foundation's Juvenile Detention Alternatives Initiative Data Analysis: Recommendation of Standards*. Olympia, WA: Washington State AOC available at <http://www.courts.wa.gov/wscrr/?fa=ccr.publications>

Summary

In March 2008, the Governor's Juvenile Justice Advisory Committee contracted with the Washington State Center for Court Research to 1) perform an evaluation, 2) make recommendations, 3) provide technical assistance, and 4) conduct analyses of the Annie E. Casey Foundation's Juvenile Detention Alternative Initiative in Washington State. Juvenile courts in King, Pierce, Spokane, Whatcom, and Benton/Franklin counties participate in the program.

This report is the third in a series, following the Data Capacity Assessment and Recommendation of Standards, scheduled for release between June 2008 and February of 2009. This report describes the ongoing implementation process used for data analysis that will demonstrate the adoption of recommendations to reconcile JDAI reporting. The final report will describe the result of efforts to implement these standards and the outcomes of JDAI data analysis.

This interim progress report summarizes the process for implementing these recommendations that will modify current analysis and reporting procedures and reconcile them to produce comparable data analysis, including Casey Foundation, statewide, and potential future analysis. The implementation plan can be described by a series of activities that fall within several stages: data collection, exploratory analysis, technical assistance, preliminary data analysis, validation, final data analysis, and process expansion. Each site is currently engaged at a different point in the implementation plan. The report includes an update on the status of progress of the sites through the process. The implementation of recommendations and the data analysis are proceeding on schedule.

comparisons across time, across the five JDAI sites; implementation will also support performance reporting that aggregates results from all five JDAI sites; the approach taken with the JDAI sites will also enable statewide reporting on juvenile detention practices. The final report in the series will summarize the results of efforts to implement these standards of data collection, analysis and reporting. It will include a detailed analysis of program outcomes and is anticipated to be completed by February 2009.

Background

The Annie E. Casey Foundation (Casey) pursues a variety of activities intended to more effectively meet the needs of today's vulnerable children and families.² Launched in 1992, JDAI is a Casey Foundation program that focuses on the detention component of juvenile justice. The objective of JDAI is to reduce the unnecessary detention of juveniles. The goals of the initiative are to:

1. Reduce the reliance on secure confinement.
2. Improve public safety.
3. Reduce racial disparities and bias.
4. Save taxpayers' dollars.
5. Stimulate overall juvenile justice reforms.

JDAI currently has a national representation of approximately 100 sites across 22 states and the District of Columbia.³

GJJAC selected JDAI as a model for best-practices outcomes and, with a grant from the Casey Foundation, King, Pierce, Spokane, and Whatcom counties began implementing JDAI in 2004. Benton/Franklin counties joined the initiative in July 2007. These counties represent approximately one-half of Washington State's youth population aged 10-17 and account for approximately one-half of

Washington's juvenile referrals.⁴ GJJAC administers JDAI in Washington State through a project coordinator. GJJAC sought an agent to evaluate and assist the data collection, analysis, and reporting for JDAI in Washington State and contracted with WSCCR in March 2008.

Study Objectives

This study aims to:

- Assess the current data capacity at each of the five sites with regard to the quality and capability of available data and accuracy of analysis and reporting.
- Recommend a common set of standards in data collection, analysis, and reporting to increase consistency across sites.
- Demonstrate the adopted standards for data collection, analysis, and reporting using data from each site.

Research Design

The JDAI data analysis project first involved an assessment of the data, analysis, and reporting from each site. During the spring of 2008, data, documentation, definitions, and calculations were collected from the JDAI sites, the JDAI statewide coordinator, the JDAI Help Desk, and Casey analysts. These materials were reviewed to determine the similarities and differences between JDAI sites, compliance with Casey Foundation reporting requirements and the quality and availability of the data to report JDAI outcomes. It was found that the details of data collection, analysis and reporting differed from site to site, but that sufficient similarities and availability would allow them to be reconciled to produce comparable reports.

The second step in the analysis project was the production of a set of recommended standards. These standards were designed to reconcile the differences in data analysis and reporting across the JDAI sites. These recommendations were created so that the

² Information available on May 20, 2008 from <http://www.aecf.org>

³ Information available on May 20, 2008 from <http://www.aecf.org/Home/MajorInitiatives/JuvenileDetentionAlternativesInitiative.aspx>

⁴ Information available on May 20, 2008 from <http://www.dshs.wa.gov/ojj/JDAI.shtml>

standards are supported by available data from each site and statewide. With the guidance of a JDAI Standards Working Group, standards were developed to resolve discrepancies in definitions between sites and produce conformity in analysis and reporting. These standards will allow direct comparison of JDAI outcomes between sites, across time and statewide, while continuing to satisfy Casey Foundation requirements.

Implementation of Recommendations

This interim progress report provides a summary of the plan for implementing the recommendations produced by the JDAI Standards Working Group. As part of the implementation plan, there are several necessary activities to ensure the reliability of the data, the accuracy of measures and calculations, adoption and adherence to the recommendations, and validation of outcomes with the sites. These activities fall within several broad stages of implementation: data collection, exploratory analysis, technical assistance, preliminary data analysis, validation, final data analysis, and process expansion.

This report outlines and describes the activities that comprise each of these stages. It is necessary for each site to participate in this process and play an vigorous role in each activity to ensure successful outcomes. Each site is currently engaged in activities at different points throughout the process of implementation. The final part of this report summarizes the progress that sites have made along these stages and the current status of standardizing report production.

Data Collection

Throughout the beginning of the project, sample data was collected from the JDAI sites as part of the assessment of capacity. Some sites delivered a full set of data including pre-JDAI historical records. Others provided a sample or subset extract. The data extracts, along with specific detailed information provided by the sites, were sufficient for assessing the

data quality and capacity of each site, as well as developing detailed reporting recommendations.

The data extracts that were previously provided are also being used to support implementation of the recommendations through the initial stages of the data analysis. However, for all sites, it will soon be necessary to provide a full extract of detention, diversion and RAI data. For those that had previously provided a full extract, it will be necessary to update the extract to include all record history to date. In the future these extracts will need to be refreshed periodically prior to reporting.

Data from the state Juvenile Court System (JCS), will come from an existing research database extract of all historic detention data to date. This database extract is repeatedly updated, and is currently awaiting modification prior to the next extract. The planned modification will better incorporate the use of aliases to more closely identify individual youth.

Exploratory Analysis

Each JDAI site records and stores data in databases built based upon different structures and formats. Exploratory analysis consists of a number of activities to better understand the database structure, format, and the composition of records. It includes such details as understanding the type, frequency, validity, use, and relationships of records and data elements.

In order to successfully implement the recommendations in a systematic, and therefore consistent and reliable way, it is necessary to extract the information common among these databases. The information is used to combine the sets of data into one dataset with a common form and structure. Once this structure is built, it will also be possible for one common set of programming to produce identically calculated measurements, commonly defined categories, and standardized reports regardless of the original data source. Provided that the unique structure and format of the original site data is well

understood, the use of one set of programming will ensure comparability of reported measures and outcomes.

Importing Data

The JDAI sites store data in a variety of database programs. These programs, though often well suited for data entry and storage, are in some instances incompatible and not necessarily well suited for analysis and reporting. Analytical programming software that is better suited for reading data from a variety of database software programs provides the flexibility to accommodate the different datasets. Furthermore, this type of software has special tools to manipulate and process the data, and reproduce calculations in routines, for repeated applications.

The goal is to have site data arrive for analysis without any modification, and in a format specific and convenient for the database program used at the site. This should benefit the sites in saved time and resources. The unprocessed data is imported into the analytical software using routines that make this process quick and automatic.

Preprocessing

A necessary step in analysis is the preparation of the unrefined data by preprocessing. Exploratory analysis often reveals details about the data that need explanation. For instance, some datasets are designed to record data as “missing” if no data are present for a particular data field, while others have a code that automatically enters a substitute, non-missing value for missing or invalid entries. Reporting these entries as valid values would produce inaccurate results.

It is often the case that data extracts require some degree of modification. An example is when variables have different formats. These formatting differences can prevent data from being grouped or calculated with similar types of data. This is often the case with different time and date formats, but occurs with other types of data as well.

In rare instances the data requires cleaning. Some databases limit entries to certain valid values, for

example, male (M), female (F), and unknown (U) for gender. A record that has a race code appear as an entry for gender may indicate that those safeguards are not present. One alternative is that this may be evidence of an error that was introduced through manual alterations to the database. Once identified, data cleaning can compensate for these situations.

Preprocessing of the data helps to prepare the data, transforming it into a more usable form for data analysis. Similar to programming for importing data, data modifications and cleaning can be programmed so that it is quick, routine, and specific to the particular data requirements from each site.

Technical Assistance

Throughout the data analysis process and with heightened importance during the exploratory analysis and preliminary data analysis stages, WSCCR will require technical assistance from sites. To properly understand the diversity of database structures and formats across all five sites would require a tremendous amount of time and resources. To make the work feasible within the limits of the project timeline, regular interaction with site staff during the following weeks and months will ensure that the representation of the records in each dataset is well understood and used properly.

The extracts of data from the sites contain a wealth of information about their structure, format and use, however they are not self documented. As WSCCR imports, reformats and cleans the data, it will be necessary to check with the sites to confirm that these actions are necessary, appropriate and reasonable. Similarly, as the programming is built and refined to calculate JDAI measures and reported outcomes, assistance and expertise from the sites will be vital.

Preliminary Data Analysis

The preliminary data analysis will focus on the reproduction of previously reported outcomes. This may seem contrary to the goals of the research project but it serves two primary purposes. The first purpose

will be the initial construction of programming for calculations such as average daily population and average length of stay. This programming will serve as the basis for subsequent generalization to accommodate data from all the sites, and refinements that will incorporate the recommended analysis and reporting standards.

The second purpose of reproducing prior reports will be to better understand the details of the analysis performed at each site. In order to develop accurate and reliable calculations, it will be necessary to reconcile the calculations between all the sites, and correct any miscalculations from any particular site.

The data capacity assessment first identified similarities and differences between reported outcomes. This assessment was performed without access to the actual calculations. The opportunity to reproduce the data will help investigate the similarities and differences of analysis between the sites, and identify where modifications are necessary to standardize calculations.

Calculated measures produced through an original development process will provide insight into possible miscalculations in previous site reports. Likewise, previous sites reports will provide guidance to correctly design accurate measures and calculations. Arriving at measurements that agree will help confirm accurate procedures. The exact reproduction of previous reported outcomes is the favored result. It would indicate that the data is used in identical ways and is processed using similar steps. This however will not always be the case, and the attempt to reconcile differences will provide equally valuable information. Differences arising during the reproduction of the data will indicate possible data errors, calculation errors or incomplete understanding of the data. This information will be helpful to work with the sites and make modifications to the programming where necessary, or explain where differences are justified and valid.

Validation

Each site will perform a critical task by validating the preliminary data analysis outcomes as they are produced and refined. Throughout the development of data analysis programming WSCCR will take interim results, contrast them with those previously produced by the sites and submit a comparison of the two outcomes for the sites' review. The site, with assistance from the WSCCR, will interpret the results and any differences in reported outcomes. Where possible, WSCCR will provide documentation describing the nature of the differences and the possible causes, or questions that will help determine which calculation is accurate. The sites will provide feedback in areas where corrections are necessary. This new information will help guide modifications needed in the data analysis to increase accuracy. Otherwise, WSCCR and the sites need to explain why the discrepancies in reporting should exist, and determine that the new programming is correct.

For efficiency this validation may occur piecemeal, with the sites validating certain components of reported measures as they are produced. Most sites can anticipate this being a recursive process. Feedback from preliminary data analysis will help refine the analytical tools, and new outcomes will need re-evaluation to continue the refinement.

Experience thus far indicates that most differences in reported outcomes appear to arise for three reasons. The first and most common reason is a difference in the criteria for selecting records used in a calculation. If the set of individual records included in calculations are slightly off, total counts and averages will disagree. They can easily be corrected through site feedback, to better understand the structure, definition and use of the data.

The second reason is due to a difference in the method of calculation. This may indicate where one set of calculations is performing incorrectly, and needs revision. The third reason is the need for different or updated data. Since the compared measures are being

produced using different means, they may also be using slightly different data. This may be due to different datasets, or data drift, where records are routinely updated over time.

Sites should anticipate that the reproduced data and the original reports may not match exactly. In the reports, prior calculations are repeatedly used in subsequent steps. For example, information from several sources is used to create measures such as average daily population and average length of stay. Disagreement in one place will result in differences elsewhere. Also, if sites used certain calculations that round time variables, calculating to the nearest hour, or day, the results will likely differ but remain approximately the same.

Prior to the release of any of the results of the preliminary analysis, WSCCR will submit a copy of the reproduced report for final approval from the site. Validation will be complete when the new analysis tools are reproducing prior reports accurately. What is important is that the results are substantially similar, or any differences are satisfactorily explained.

Final Data Analysis

Once the validation of the data is complete, the next stage will be to prepare for producing a finished report product. WSCCR's production of results substantially similar to that produced by the sites will serve as confirmation the data collected is sufficient, the calculations are accurate and the outcomes measured are the results of interest. The next step will be to generalize the site specific formatting and calculations. That process will incorporate the recommendations for standardizing the analysis. The final product will be a common set of programs that produces the recommended measures and outcomes.

The recommendations produced by the JDAI Standards Working Group and released in the previous publication are specifically designed to resolve some of the inter-site differences in analysis and reporting. Incorporating some of the recommendations will not be labor intensive, and in fact, some of the work can

be done in the early stages of importation, data cleaning and preliminary analysis. Much of the work is re-categorization and regrouping. In some cases the calculations are already developed and only require the uniform and consistent application to data from all of the sites.

Other elements of the recommendations may require considerable attention, such as creating and using a unified law table. As noted in the previous reports, it is also likely that throughout this process additional challenges and incompatibilities may arise and need resolution.

After the recommendations are incorporated into the data analysis, both the final report outcomes and the report format will be different than that previously produced by the sites. This is intentional and expected. The standards recommended differed in part from the reporting practices at every site. Even where data previously agreed, the different categorizations and presentation in reports will make results appear different.

Process Expansion

One of the products of successfully incorporating the recommendations into data analysis will be a single set of programming code that performs the calculations used for JDAI reporting, regardless of the data set to which they are applied. This creates new opportunities for the creation, use and distribution of the data.

One opportunity to improve the new analysis, and one which will not have a great deal of visibility once it is complete, is the automation of analysis and reporting practices. As the existing recommendations are incorporated, new recommendations, reporting formats, and measures of interest are likely to be developed to deepen the data analysis and answer new and interesting questions. Automated analysis can still be modified in the future. It is important to maintain flexibility to adapt to changes in focus, presentation of data, reporting methods, and source data. However, analysis that is seen as important and

that is likely to be repeated in the future should be automated to save time and resources.

When it becomes available, it will be necessary to perform a data analysis process on statewide data similar to that outlined above for site data. Having already produced much of the original work for the sites, only minor modifications should be necessary to incorporate statewide data.

Standardization of analysis and reporting will build a foundation for future expansion of reported measures and outcomes. Since the analysis process includes original site data, there is the opportunity for the deepest possible analysis. One likely improvement will be the inclusion of annual summary reports into automated routines for each site and across Washington State. It will be possible to more efficiently and effectively investigate site trends and compare outcomes of programs within sites, between sites, statewide and across time.

Finally, pieces will be in place to increase access to the data and the reports in the future, thereby unburdening sites with the demands that repetitive calculations and report production require.

Current Project Status

The five JDAI sites are currently engaged at different points throughout the data analysis process. Site specific challenges are encountered at each stage and will be dealt with as they arise. However, each site will proceed through the stages in a similar fashion, and development will proceed in parallel.

All sites had previously provided data extracts and those datasets have been successfully imported for exploratory analysis. WSCCR is currently collecting statewide data, and shortly sites will receive a request for an update to their detention, alternative and RAI data. Again, for those sites that have not already done so, this updated extract must include the full record history to date. Programming is under development and for most of the sites the programming to automate importation and reformatting of the data is complete.

As analysis of site data proceeds through the exploratory stage, WSCCR has begun to contact several sites requesting additional information and clarification on the use and representation of the data. The sites that are not currently providing technical assistance can expect to be contacted shortly as attention focuses on their data and questions arise in the reproduction of reported measures.

Preliminary data analysis has been conducted for two of the sites with regard to demographic information and is just underway for a third. Results indicate a strong need for validation of this preliminary analysis; however the outcomes produced by WSCCR are substantially similar to site results and in some cases identical.

Perhaps most significant, WSCCR has successfully developed preliminary programming for extracting population, detention alternative, and RAI statistics including average length of stay and average daily population. This will likely speed the development process for other sites, and serve as the foundation for building the recommended reports. For those sites that have progressed to this stage, these calculations appear to be validated when they are compared with site reports. In addition, a comparison of the RAI and override data for the first of five sites appears to validate the work.

Next Steps

The next step in this project will be the continued collaboration with sites to implement these recommendations. After the measures are refined and the results are validated, the programs will be generalized, combined and then modified to satisfy both Casey Foundation requirements and the working group recommendations. Statewide data will soon be available and this too will be combined with the site data.

Afterwards, it will be necessary to develop the reporting templates within the analytical software. These are the recommended reporting templates designed to standardize reporting for the Casey

Foundation and statewide, and can be found attached to the Recommendation of Standards report. This will seamlessly integrate the data analysis procedures with report production.

Eventually, WSSCR will establish a data sharing relationship with the sites and continue to periodically receive unprocessed data extracts. The programming developed for importation, reformatting, analysis and reporting will create the foundation for a centralized “data mart” for JDAI data.

Final Report

The final progress and outcomes report will be a more detailed assessment of the research project. It will summarize the later stage progress towards adopting uniform standards, verify the reliability and accuracy of analysis and reporting, and compare JDAI sites and statewide data. The final report is anticipated to be completed by February 2009.

For further information, contact: Edward Valachovic at (360) 705-5336 or
Edward.Valachovic@courts.wa.gov

WSSCR is the research arm of the AOC and was established in 2004 by order of the Washington State Supreme Court. The WSSCR conducts research to improve the understanding of the courts, help guide judicial policy, and improve the functioning of the judicial system.