Enhancing Child Support Enforcement Efforts: Summary of Data Warehouse Efforts in Nine States

Final Report

Prepared for:

Department of Health and Human Services

Office of the Assistant Secretary for Planning and Evaluation/Office of Human Services Policy and the Administration for Children and Families/Office of Child Support Enforcement

Prepared by:

The Lewin Group

And its subcontractor:

SRA International, Inc.

October 2006

Enhancing Child Support Enforcement Efforts: Summary of Data Warehouse Efforts in Nine States

Final Report

Prepared for:

The Department of Health and Human Services

The Office of the Assistant Secretary for Planning and Evaluation/Office of Human Services Policy and the Administration for Children and Families/Office of Child Support Enforcement

Prepared by

The Lewin Group:
Karen Gardiner
Mike Fishman
Asaph Glosser
Matthew Langley

SRA International, Inc.:
David Vennergrund

This report was produced under the direction of Linda Mellgren (ASPE) and Dennis Putze (ACF/OCSE) under ASPE contract 282-98-0016, Task Order 32, Jerry Regier, Principal Deputy/Assistant Secretary for Planning and Evaluation. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the federal agencies that funded or reviewed this report. This report is available on line at: http://aspe.hhs.gov/hsp/07/CSE-enhancement/data/index.htm

Table of Contents

I. INTRODUCTION

A. Project background

State automated child support systems support a number of functions. They must contain all data necessary to manage cases and meet Federal reporting requirements.[1] There are case initiation functions (e.g., accept automated referrals from the welfare, foster care, and Medicaid agencies and other state child support programs and maintain case records); locate functions (e.g., interface electronically with state and Federal sources to obtain and verify locate, asset, and other information about parents); establishment functions (e.g., track, monitor, and report the status of paternity and order establishment); case management functions (e.g., automatically update cases and provide information to other programs, support review and adjustment of orders); enforcement functions (e.g., monitor compliance with support orders and initiate enforcement actions such as income withholding and tax refund offsets); and financial management functions (e.g., process payments received, disburse payments).

Statewide systems also support Federal reporting. The 1996 Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) includes a provision requiring state Child Support Enforcement (CSE) programs to conduct annual self-assessments of their performance in eight areas[2] and submit their findings to the Department of Health and Human Services (HHS)/ Administration for Children and Families (ACF)/ Office of Child Support Enforcement (OCSE). The self-assessment process allows states to manage their CSE programs and explore ways to improve performance. Administrative data are also important in the context of the incentives OCSE pays to state CSE programs. Since 1975, the Federal government has paid incentives to encourage improved collections through efficient establishment and enforcement techniques. These incentive payments are a key source of funding for many state CSE programs. The method for calculating incentive payments changed with the adoption of the Child Support Performance and Incentive Act (CSPIA) in 1998. Incentive payments are based on performance in five areas deemed by a Federal working group to be most critical to supporting families.[3]

Statewide child support automated systems contain a wealth of valuable data that is used for program operations, and could also be used for measuring performance, improving data quality, supporting decision-making, and responding to inquiries from state officials or other interested parties.[4] The nature of the systems in many states, however, makes it difficult to extract and analyze data in a timely and cost efficient manner. Thus, states have begun to use data warehouses[5] and other tools that use data extracted from statewide systems to create reports and other statistical information for program managers and staff. States use the information

---

1 Office of Child Support Enforcement (OCSE) (2000). *Automated Systems for Child Support Enforcement: A Guide for States.* Washington, D.C.: Administration for Children and Families, U.S. Department of Health and Human Services.

2 Case closure, establishment of paternity and support orders, enforcement of orders, disbursement of collections, securing and enforcing medical support orders, review and adjustment of orders, interstate services, and expedited processes.

3 Paternity establishment, order establishment, collections on current support due, cases paying toward past-due support, and cost-effectiveness.

4 OCSE Dear Colleague Letter DCL-01-29, June 15, 2001. Available on-line at http://acf.dhhs./gov/programs/cse/pol/dcl-01-29.htm.

5 While states vary in their definitions, a data warehouse generally comprises a computing system that stores information about an organization's activities in a database. The database design facilitates reporting and analyzing data in order to obtain information and facilitate decision making. See http://en.wikipedia.org/wiki/data_data warehouse for more information.

generated from the data warehouses for Federal reporting, for policy development, and monitoring the performance and improvements of local offices. These data warehouse are subject to the same security and privacy protections that cover other child support uses at the state level.

The Department of Health and Human Services /Office of the Assistant Secretary for Planning and Evaluation (ASPE)/Office of Human Services Policy (HSP), in collaboration with OCSE, contracted with The Lewin Group and its subcontractor, SRA, to learn more about state experiences with data warehouses; that is, how they are used to help manage program operations, assess program effectiveness and efficiency, and inform policy development. This report summarizes findings from discussions with staff in nine states about the development, content, and capabilities of their data warehouses.

B. Methodology

Three sources informed the selection of study states. First, we reviewed the summary of states that participated in OCSE-sponsored conference calls that focused on using child support data to support performance improvement and decision making. OCSE facilitated five calls with states in the summer of 2001 to encourage information sharing about data warehousing, data mining, and other techniques. Second, we reviewed the plans of the four states that received 1115 grants from OCSE to construct or enhance data warehouses. Finally, the Federal project officers suggested a number of states that had a history of using data for research and program management.

In all, we had 14 candidate states. We had brief conversations with each state to learn:

- Data included in the data warehouse
- Data uses (e.g., program management, policy development)
- Security and privacy safeguards

Based on the calls, we proposed nine states for inclusion in the study. We conducted site visits to five states and had telephone discussions with four states:

Site Visit States	Telephone States
Maryland	Maine
Michigan	North Carolina
Pennsylvania	Wisconsin
Vermont	Wyoming
Washington	

The site visits and calls occurred in winter and spring 2004. Prior to the conversations, we developed a discussion guide that explored five specific areas: data warehouse background, data sources, data uses, data users, data safeguards, and funding sources. We used the same guide for site visit and telephone states. In the site visit states, staff also demonstrated the data warehouses' capabilities.

C. Organization of Report

The report is divided into the following sections:

Overview of findings. The first section summarizes the findings from the discussions across a number of dimensions, including primary uses of the data warehouse, data sources, types of reports produced, and ad hoc reporting capabilities. This section also summarizes respondent comments on lessons learned during development and implementation of their data warehouses.

State case studies. Each case study describes the following:

- *Data warehouse background*. We asked respondents when the data warehouse was created, the impetus for creating it, and key individuals involved in the decision to create it.

- *Implementation issues*. The respondents described the planned timeline for the data warehouse development, any obstacles to implementing the data warehouse, what (if anything) needed to be changed, and whether, in hindsight, they would alter anything about the data warehouse design.

- *Data sources*. The respondents described the sources of the data in the data warehouse, the level of detail, how it was determined which elements would be included, and whether there is capacity (and if so, plans) to add new data elements.

- *Data uses*. The respondents described the primary and secondary uses of the data warehouse, including reports that are produced regularly (e.g., Federal reports) and ad hoc capabilities. We also explored the type of analyses conducted using the data warehouse data. With regard to Federal reports, we asked about the following capabilities: We also identified data security and privacy protections, including safeguards for users access to the data.

    - Federal incentive-related reports. We asked if the state staff use the data warehouse to create the reports used to calculate Federal incentive payments to the states. There are three key incentive-related reports: the OCSE 157, the OCSE 34A, and the OCSE 396A. The data reported by states on the OCSE 157 are used to calculate four of the five performance measures—paternity establishment, cases with orders, current collections, and cases paying towards arrears. The OCSE 396A (Child Support Enforcement Program Financial Report) and OCSE 34A (Child Support Enforcement Program Quarterly Report of Collections) are used to calculate the cost-effectiveness measure. The OCSE 34A also is used to determine state collections.

    - Self-assessment reports. We also asked whether state staff use the data warehouse to produce self-assessment reports. As noted above, PRWORA requires each state to annually assess the performance of its child support enforcement program in eight areas and to provide a report of the findings to the OCSE Commissioner. The self-assessment focuses on whether child support service delivery is in accordance with Federal mandates. There are no financial sanctions associated with the self-assessment; rather, it is a management tool for states to evaluate their performance. OCSE may provide states with comments, recommendations for corrective actions, or technical assistance.[6]

[6] For additional information about the self-assessment, see Action Transmittal OCSE-AT-00-09, Final Rule on State Self-Assessment. Available on-line at http://www.acf.hhs.gov/programs/cse/pol/AT/at-00-09.htm.

- *Data warehouse users*. We explored which child support staff (e.g., state policy staff, county office managers, caseworkers) have access to the data warehouse. Because some of the data warehouses use data from a number of different programs, we also identify those states where non-child support staff (e.g., staff from other state agencies or universities) can access the data warehouse.

- *Funding*. We asked respondents about the primary funding sources of the data warehouse, the initial start-up costs, and maintenance costs.

- *Architecture*. A matrix summarizes the technical details of the data warehouse.

II. OVERVIEW OF FINDINGS

There is no "one size fits all" data warehouse. As the case studies (**Section III**) indicate, states use their data warehouses for a variety of purposes. Some focus on Federal reporting; others use the data warehouse for specified child support activities (e.g., locate). Some use the data warehouse to manage their programs (e.g., identifying underperforming offices, overstaffed functions), while others use it to explore and assess new policies. The data warehouse data come from a variety of sources. Each state's automated statewide system plays a prominent role. Also, in some states, the data warehouse directly interfaces or feeds into other systems (e.g., welfare). State data warehouses also vary in terms of reporting capabilities. Some use them to generate Federal reports for the incentive-related data reliability audits or the annual self-assessments. Others use their data warehouses primarily for ad-hoc reports. State child support programs also grant different levels of access to their data warehouses. For example, in some instances, partner agencies can access limited data; in others, access is highly restricted. Finally, the state respondents had unique perspectives on best practices and advice for other child support programs interested in building data warehouses. Each of these areas is discussed further below.

A. Data warehouse Uses

We asked respondents why their states invested in a data warehouse. They cited myriad primary and secondary data warehouse uses (see **Exhibit II.1**), but most noted that a key data warehouse benefit, in comparison to a standard transactional database, is its ability to speed up queries.

Exhibit II.1: Data warehouse Uses

State	Primary Use	Secondary Use
Maine	Program management	Policy development
Maryland	Policy development	Federal reporting
Michigan	Parent locating	Federal reporting
North Carolina	Federal reporting	Program management
Pennsylvania	Federal reporting	Program management
Vermont	Program management	Policy development
Washington	Policy development	Program management
Wisconsin	Policy development	Program management
Wyoming	Medical support enforcement	Program management

Respondents in two states indicated that the data warehouses were created primarily to help manage the child support program or to create Federal reports. For example:

- Maine's data warehouse is used primarily for program management. Prior to creation of the data warehouse, staff had to rely on the statewide automated child support system to assess performance of specific caseloads (e.g., cases in the intake function) or districts (e.g., the percent of cases with orders in a given region of the state). The data warehouse negated the need to create COBOL programs to track performance. State-level staff use predefined reports to monitor local-level staff performance and allocate staff resources as necessary. Local-level staff can also use the data warehouse to monitor their caseloads.

- Vermont also uses its data warehouses primarily for program management. Vermont's IV-D director stated that the goal of the data warehouse was to increase productivity by staff and program effectiveness by enabling users to easily convert data into information that is accessible and understandable, which will in turn result in an "actionable" activity. Examples of reports that lead to actionable activities include cases to review for closure, cases with arrears not receiving payments, and arrears-only cases with low arrearage amounts (indicating a need to contact both parties and prompt a settlement).

Staff in three states indicated that their data warehouses were used primarily for policy development-related research.

- Washington's data warehouse is used to conduct policy development research and statistical analyses that help staff to understand the consequences of the child support agency's actions and policies on outcomes for the agency and customers. For example, it is used to create statistical models that facilitate cause and effect analysis (e.g., how caseworker actions affect outcomes) and cost avoidance studies (i.e., how regular child support payments produce savings to the welfare, Medicaid, and Food Stamp programs). Staff use the data warehouse to learn about causal relationships between individual characteristics, timing of life events, and the payment of child support. For example, the data warehouse will be used to analyze order amounts to understand the consequences of default orders and imputed income. State staff will also explore hard-to-collect cases and what collection techniques tend to work.

- Wisconsin staff use the data warehouse to efficiently access and analyze aggregate data to better understand the interrelationships among state programs. For example, it is used to track earnings and Temporary Assistance for Needy Families (TANF) benefit receipts for TANF recipients prior to and following participation in employment programs. Findings are used to make a business case for operating programs.

- Maryland's data warehouse supports policy development. For example, one study found that 80 percent of welfare customers in Baltimore failed to attend their child support intake interviews. The study's findings were used to develop a pilot intervention. In another example, staff used the data warehouse to study patterns of benefit receipts among children born to unmarried parents for whom an in-hospital paternity acknowledgment had been completed. The study found that within one year of birth, 33 percent of children whose paternity is acknowledged become known to the child support enforcement agency and 75 percent to the Department of Human Resources (the umbrella agency for TANF, child support, Food Stamps, and Medicaid). The research informed cross-program planning between child support, TANF, and other programs that serve low-income families.

Other states use their data warehouses for specific program functions.

- Michigan's data warehouse is used primarily for locate operations. The data warehouse integrates child support data with elements from other state agencies and external sources (e.g., financial institutions). CSE staff use it to locate addresses for parents and new employers.

- Wyoming's data warehouse is used to track and enforce medical support orders. It provides a clearer picture of medical support in the state, including the number of children with insurance, the number of children receiving public health benefits (e.g., SCHIP), and the number of employers offering medical insurance benefits. Staff will be able to compare these statistics to actual court orders for specific cases (e.g., the child is receiving coverage as ordered).

Finally, two states (Pennsylvania and North Carolina) primarily use their data warehouses to create Federal reports (see **Section C** for more detail).

B. Data Sources

Staff in the nine states indicated that data sources and specific data elements are determined by the data warehouses' proposed uses. For example, a data warehouse that is created to support Federal reporting (e.g., Pennsylvania) will contain different data elements than a data warehouse designed to conduct cross-program policy assessments (e.g., Wisconsin). Staff in a number of the study states referenced user groups or advisory committees that met in the early stages of the data warehouse's development to determine the specific data elements that would support the primary function(s). All respondents indicated that additional data could be added as needs arise or additional funds become available.

The primary data source for most of the data warehouses is the statewide automated child support system (see **Exhibit II.2**). In three states—Maine, Pennsylvania, and Wyoming—the child support statewide system was the *sole* data source. In each state, this source interfaced with other data systems. For example,

- The Pennsylvania data warehouse source is PACSES, the automated child support system. PACSES interfaces with systems from TANF, Unemployment Compensation, Department of Revenue, the Department of Motor Vehicles, the credit bureau, the State and Federal Directories of New Hires, the state disbursement unit and the professional license bureau. The data in the data warehouse supports creation of Federal reports (both incentive-related and self-assessment reports), as well as internal enforcement reports (e.g., Drivers license suspension report, Credit Bureau report, collections reports).

In six states, the data warehouses include data from the child support system and other sources. For example:

- Maryland's data warehouse includes data from the Client Information System (the umbrella system for the child support, TANF, Medicaid, and Food Stamps programs), as well as data from the Unemployment Insurance system (wages), the Maryland Directory of New Hires, the child care program, the Work Opportunities program (the state welfare-to-work program), and the in-hospital paternity establishment program. The data warehouse supports child support program and cross-program policy development. Staff, for example, use the data warehouse to conduct a longitudinal study of welfare leavers, including demographic characteristics, welfare recidivism, and post-exit employment patterns.

- Michigan's data warehouse includes data from 12 state programs, including the child support automated system, the state Department of Treasury, the Department of Corrections, the Department of Natural Resources, the State Directory of New Hires, State Police, the Department of State, Bureau of Workers and Unemployment Compensation, and the Department of Community Health. The data support locating absent parents and employers.

Exhibit II.2: Data Warehouse Sources

State	Primary Source	Examples of Other Sources (included in data warehouse or via interface)
Maine	Child support automated system	Child support system interfaces with TANF; Department of Labor; Foster Care; Department of Revenue; Workers Compensation
Maryland	Umbrella system for child support; TANF; Medicaid; Food Stamps programs	Data warehouse includes data from State Directory of New Hires; child care program data; welfare-to-work program data; in-hospital paternities; Unemployment Insurance system
Michigan	Child support automated system	Data warehouse includes data from 12 state program systems
North Carolina	Child support automated system; TANF; Medicaid; Food Stamps	Data warehouse includes data from mental health and child development programs
Pennsylvania	Child support automated system	Child support system interfaces with TANF; Unemployment Compensation; Department of Revenue; Department of Motor Vehicles; credit bureau; State and Federal New Hires Directories; professional license bureau
Vermont	Child support automated system	Data warehouse includes data from the state financial system and Unemployment Insurance system
Washington	Child support automated system	Data warehouse includes data from the Employment Security Department (wages) and the TANF data warehouse
Wisconsin	Child support automated system	Data warehouse includes data from the Unemployment Insurance system; umbrella system for TANF, Medicaid, Food Stamps programs
Wyoming	Child support automated system	The child support system interfaces with TANF; State and Federal Directories of New Hires; state Department of Labor; Department of Motor Vehicles; Workers Compensation; Bureau of Game and Fish; Social Security Administration; Medicaid; Blue Cross (operator of SCHIP program)

C. Reports Generated by Data warehouses

As **Exhibit II.3** indicates, staff in five states use their data warehouses for Federal reporting. Four (Maryland, Michigan, North Carolina, Pennsylvania) use the data warehouse to create the annual OCSE 157 report. Three also use the data warehouse to produce the quarterly OCSE 34a report. And, three states produce the annual self-assessment reports with the data warehouse (Vermont, in addition to North Carolina and Pennsylvania). Two additional states (Maine and

Washington) use the data warehouse to select the self-assessment sample, but not to create the actual self-assessment reports.

Exhibit II.3: Data warehouse Reports

State	Federal Reports	Other Reports
Maine	None	Predefined performance and caseload reports; canned queries
Maryland	OCSE 157, OCSE 34a	Drivers license suspension; paternity affidavit
Michigan	OCSE 157, OCSE 34a	Caseload report
North Carolina	OCSE 157; self-assessment reports	60 child support reports for program management
Pennsylvania	OCSE 157, OCSE 34a, self-assessment reports	Drivers license suspension, new hires collections, Financial Institution Data Match report, credit bureau report
Vermont	Self-assessment reports	Pre-defined reports include cases to review for closure, number of actions taken by workers, cases with arrears not receiving payments; Digital Dashboard
Washington	None	Progress reports on state and Federal performance measures
Wisconsin	None	Ad hoc reports; users can create summary level data
Wyoming	None	20 predefined reports

Michigan uses Federal reports for additional purposes. State-level staff use data from the OCSE 157 and 34a reports to track local office performance. The state distributes performance-based incentives to individual counties based on these reports. Local offices use the reports to track their own performance on the five incentive measures. And, the state uses the Federal reports, in addition to a caseload report, to administer contracts for the local child support offices.

The respondents in these five states indicated that the data warehouses are also used to create a host of other reports. For example some states create enforcement-related reports.

- Pennsylvania creates four "enforcement remedies" reports: the drivers license suspension report (i.e., payments attributable to the license suspension program), the new hires collection report, the Financial Institution Data Match report, and the Credit Bureau report (i.e., the number of defendants eligible for credit reporting). A monthly report card for management-level staff summarizes collections by month and method (e.g., wage attachment) and history (i.e., the state's performance during the same period the previous year).

- Maryland staff use the data warehouse to create monthly statistics reports on selected topics. These include reports on the driver's license suspension program and the in-hospital paternity affidavit program.

Some states use their data warehouses to produce caseload management reports. For example:

- Vermont's data warehouse produces reports such as: cases to review for closure, potential audit error case list, children in need of paternity action, number of actions taken by workers, and arrears cases with low arrears amounts. A "Digital Dashboard" provides summary information on performance, such as monthly collections compared with the state's current year collection goal and prior year collections.

- Maine state-level and regional managers and supervisors use predefined reports on performance and caseloads to monitor staff performance. Examples of reports include number of orders established, non-paying cases that should be transferred to another unit, cases closed, and cases where all children are emancipated.

- Washington's staff produce monthly progress reports on Federal and state performance measures that are available at the state level, office level, and individual worker level. For each report, staff can examine the monthly goal versus actual performance.

In addition to standard reports that are available, data warehouse users in eight states can create ad hoc reports (not shown). In some states, data warehouse users can create customized reports using "canned" or predefined queries. In others, users can create their own reports by selecting from a number of attributes or "data clusters." For example:

- In Maine, canned queries are specific to an individual worker's caseload and can be used to monitor or assess performance (e.g., cases that meet license revocation criteria).

- In Pennsylvania, users can sort and filter data through "cubes". Cube subject areas include collections, disbursements, distributions, obligations, and support orders. The user can filter information across a number of dimensions (e.g., geography, date, case category, TANF status) and customize his or her report.

- Staff in North Carolina can run their own queries monthly when the data warehouse data is updated. Queries are used for case clean-up and performance tracking.

- Washington users can formulate ad hoc queries at the state, field office, team, or case worker level (e.g., a user could determine how many of a team's cases had no order in place).

- Vermont users can extract data using an on-line processing tool and manipulate it through graphs, tables, and matrices. A user could select from case data, case financial data (e.g., collections), and person data (e.g., custodial parent) at the state or local level and explore aspects of cases in different parts of the state.

In other states, ad hoc requests must be forwarded to a programmer.

- Maryland data warehouse staff received about 500 ad hoc report requests from state- and local-level child support officials annually. Generally, the reports answer case management questions, such as how many cases in a given office are in the locate function for a specified amount of time. Local offices might request a report on the number of cases not paying or

the number of cases in arrears. Managers will give their staff information related to cases to take action as appropriate.

- In Michigan, staff can request ad hoc reports through a help desk. Requests include case information (e.g., list of children in a region without paternity established, summation of payers in arrears who are incarcerated or deceased) or they can be used to track or assess new practices. One county, for example, might suspend deer hunting licenses in September to increase collections.

D. Data Warehouse Users and Safeguards

State child support enforcement agencies have access to many powerful databases that provide timely and accurate data, which in turn improve their services to children and families. Security and privacy of these data are important issues for all those involved in the child support program. We asked respondents to describe who had access to the data warehouse and whether all users had access to the same information. We explored whether individuals outside of the child support agency (e.g., staff in other state agencies, universities) had access to the data warehouse and if so, the level of access. We asked respondents what their procedures were regarding securing and safeguarding the data available through their data warehouse. The security requirements for each state are shown in the Architecture section of the case studies. The states use various means to safeguard the data: States limit who has access to the data; limit what data individuals have access to; and require users to sign non-disclosure agreements. Firewalls and passwords are used to prevent unauthorized access. See cases studies for specific security requirements by state.

As **Exhibit II.4** indicates, five states grant some level of access to all child support staff. In two of these states, all CSE staff have access to the same level of detail.

- In Vermont case workers and supervisors have access to the same screens and can run ad hoc reports. Although access is universal for child support staff, the agency anticipated there would be three types of users: power users (i.e., users who would create many ad hoc queries), moderate users (i.e., those who would primarily use pre-defined reports, perhaps adding their own fields), and casual users (i.e., those who would view summary information on the digital dashboard and some pre-defined reports).

In the other three states, only technical staff and/or managers have access to more detailed data (e.g., individual level data) or the ability to create reports.

In the remaining four states, access to the data warehouse is more restricted. In some states (e.g., Pennsylvania) state-level staff have access to the data warehouse but only limited county-level staff do so. In Maryland, child support agency staff do not have access to the data warehouse. The University of Maryland-Baltimore houses the data warehouse and staff in the Family Welfare Research and Training Group run queries and conduct studies on behalf of CSE staff. The Wyoming data warehouse will have 24 users, including state-level staff, managers from each of the nine districts, and one court official.

Exhibit II.4: Data warehouse Users

State	Child Support Users	Other Users	Type of Data Access
Maine	All child support staff	None	Child support line staff have access to predefined reports and queries; supervisors and state managers have access to individual-level data and ad hoc capabilities
Maryland	None at this time	Four programmers, three study directors, research analysts at the University of Maryland-Baltimore County	Programmers and study directors can access data with identifiers; others use de-identified data
Michigan	All child support staff	Department of Community Health (DCH), Office of the Inspector General (OIG) and Bureau of Workers and Unemployment Compensation (BWUC)	Child support staff have unrestricted query access; DCH can view health data, OIG and BWUC have access to data for fraud detection
North Carolina	All child support staff	Other Department of Social Service (DSS) programs (TANF, Food Stamps, Medicaid); Other Department of Health and Human Services (DHHS) divisions: Mental Health, Child Development, Public Health, Aging	All child support users can access canned queries and write their own reports; other DSS staff and DHHS divisions can only view data specific to their programs
Pennsylvania	State-level staff and several users in each of the 67 counties	None	Access to data is more limited for county staff
Vermont	All staff	None	Case workers and supervisors have access to the same level of data
Washington	Management and Audit Program Statistics (MAPS) staff; child support supervisors and case workers	Some tribal IV-D offices and prosecutors	MAPS staff have direct access to the data warehouse and data with identifiers; other child support staff access to de-identified data through an application; tribal IV-D offices and prosecutors who work child support cases have view-only access

State	Child Support Users	Other Users	Type of Data Access
Wisconsin	Limited staff from the child support agency	Limited staff from the Bureau of Workforce Information and the three staff from the University of Wisconsin Institute for Research on Poverty	Three levels of access: viewer (can see data but not create reports); editor (cannot create new reports); experts (can create reports)
Wyoming	Nine district managers, "super users" at the state-level, limited state-level management staff	One court official	Super users have access to reports and ad hoc capabilities; district managers and court official have access to predefined reports specific to districts/caseload

Most often, staff from other state agencies or other entities (e.g., universities) do not have access to child support data through the data warehouses. Two states grant access to university staff. As noted above, staff at the University of Maryland-Baltimore maintain the data warehouse and conduct data analyses on behalf of the Department of Human Resources. In addition, limited staff from the University of Wisconsin have access to the Wisconsin data warehouse.

In two states (North Carolina and Pennsylvania) the child support data resides in the same data warehouse as data from other programs (e.g., welfare). In both states, data access is limited to program staff (i.e., only child support staff can view child support data).

E. Lessons Learned

Respondents were asked what advice they would give to their counterparts in other states who are interested in developing a data warehouse. Each respondent was strongly supportive of the design and attributes of his or her state's data warehouse. However, all had feedback on steps that could improve the design and implementation processes.

Some pointed to the importance of building support for the endeavor. Others focused on data warehouse uses (i.e., it is crucial to determine the primary purpose of the data warehouse from the outset because it will drive the data needed). Still others described more technical issues, such as building in time to assess data warehouse capabilities as they are added and ensuring that the technical team that builds the data warehouse understands the complexities of the child support program and the agency's needs.

1. Build Support

Secure buy-in of management. A number of respondents stated that senior management sponsorship and commitment are crucial to the success of a data warehouse. Management support is not only important in terms of procuring resources; they can ensure that key staff have the time to develop the project. One respondent who provides research and statistics to the child support director noted that it was extremely helpful to have time blocks available to work without interruption; data requests (e.g., "I need this in five minutes") were kept to a minimum.

Respondents added that because senior management might change frequently, it is important to reeducate people about the data warehouse and its uses.

Involve other key stakeholders. Respondents also noted that it is important to build relationships with other stakeholders, such as staff or managers in other agencies that might provide data, prior to building the data warehouse. Creating an environment of trust is important. Meetings with counterpart agencies are an opportunity not only to discuss the types of data that the child support agency hopes to include in the data warehouse, but to explain why the data is important and how the information produced by the data warehouse could be useful to all agencies involved. One respondent stated that her agency receives data from the social services agency because senior staff see the value in the research produced, adding that the reports "make the agency look good in front of the governor and the assembly."

2. Specify Uses and Users

Clearly define goals. Respondents stated that prior to developing a data warehouse, the child support agency needs to clearly enumerate its key goals. Without a clear set of goals, the data warehouse may not attain the agency's objectives within the budget allotted. The goals inform the business and design requirements. If the data warehouse is needed to produce "X" then it should not be designed to produce "Y". It is much more difficult to change the data warehouse once development has begun. Once the key goals are established, the agency should develop, in writing, a clear understanding between the agency and the information technology providers regarding the initial goals, objectives, and costs.

Obtain input from future users. It is important to know the needs of the people who will be using the data warehouse. They, in turn, need to be confident they know what the final product will be and how it can be used. Additionally, the "views" of the information must be the ones that users want (e.g., reports, graphs, ad hoc capabilities). A number of respondents stressed the importance of developing the data warehouse with the end user in mind. Thus, it is important to include them in development discussions. The designers should know the questions that users want answered. One respondent noted that sessions with child support staff and other potential users were extremely valuable because they indicated data that might be needed in the future. The respondent added, "It is better to start with too many variables and whittle them down than to have to add them after the fact." It is also important to document the discussion with the users. Another respondent added that the data warehouse design and development group should seek the perspectives from a wide range of potential users. For instance, local-level users might have different needs than state-level staff.

Stress security and confidentiality. Data warehouses make it easier for child support staff to abstract and analyze data from numerous databases, but also can raise concerns about confidentiality. Respondents indicated that putting safeguards in place to secure data was of paramount concern. Examples of security procedures include limiting access to identifiable data (e.g., program managers and system operators), requiring users to sign a non-disclosure agreement or go through clearance that indicates level of access, limiting data access by specific role or organization (i.e., warehouses that contain data from multiple programs limit access to child support data to child support staff), and providing view-only access or access to predefined reports to non-management staff.

3. Keep in Mind Technical Issues

Develop data warehouse incrementally. Some respondents advocated developing the data warehouse in distinct iterations, adding data, a specific function (e.g., ad hoc capabilities), or both in each phase. An iterative process builds in time for data warehouse developers and other stakeholders to pause and review the work to date to determine if changes are necessary. For example, one respondent suggested that staff involved in the creation of the data warehouse let the business requirements "sit" and then revisit them before beginning the development process, thus allowing staff to reflect about what they want in the long run. In addition, an iterative approach to development provides users with access to the data warehouse's capabilities to meet some of their information needs while allowing the necessary time to develop a comprehensive data warehouse with full functionality.

If vendors are used, they should be familiar with the child support program. A number of respondents stressed that potential vendors should not underestimate the complexity of the child support program, even in small states. One noted that although her state has a small caseload, the child support program operates under the same rules and requirements as states with large caseloads. Another respondent noted that vendors should not write proposals for data warehouses. If the child support agency wants the data warehouse, its staff should write the proposal. People at the ground level who understand the program need to be involved.[7]

Respondents raised a number of other issues related to vendors. Several stressed the importance of the vendor being on site. One child support agency used an out-of-state vendor to build the data warehouse. Agency staff thought they and the vendor had an agreement that the data warehouse builders would spend a large amount of time on site. The agency's definition of "on site" (weekdays between 9:00 a.m. and 5:00 p.m.) and the vendor's (time on weekends) did not match. It is important to clarify these roles early in the process.

Others suggested that the vendor should have experience building data warehouses. The same vendor should be used throughout the course of the contract.

Finally, one respondent suggested it is useful to have an independent IT advisor to the data warehouse project who is precluded from bidding on its development and construction.

[7] Note that OCSE systems policy (OCSE AT 09-11) requires that contractors who assist the state during development of specifications, requirements or statements of work are excluded from competing for subsequent phases of the project or being included as a subcontractor. This is to ensure objective contractor performance and eliminate the appearance of unfair competitive advantage.

CASE STUDIES

MAINE CASE STUDY

Data Warehouse Snapshot

Date created	1997
Primary uses	Program management and policy development
Primary data sources	Child support automated system (NECSES)
Types of reports	Ad hoc
Users	Statewide child support staff
Primary funding source(s)	Child support

Background

Maine's child support data warehouse became operational in February 1997. It was created to facilitate program management and policy development-related research. Prior to the data warehouse's creation, the Division of Support Enforcement and Recovery (DSER) managers could not examine the performance of specific caseloads (e.g., cases in intake) or districts. They could examine only individual cases. Managers could not explore performance indicators or projections. Any analysis conducted was cumbersome—staff had to create Cobolt programs—making it difficult to identify and address issues as they arose.

The eight participants of the DSER Data User Group were actively involved in the decision to create the data warehouse. In addition, a systems analyst from the Bureau of Information Services (BIS) was involved. BIS maintains the mainframe computer, from which the data warehouse is populated, as well as the data warehouse.

The obstacles that had to be overcome were program staff time and lack of knowledge regarding data warehousing.

Implementation

The time line from planning to development was six months. There were no unexpected obstacles in implementing the design.

During implementation, some changes were made, however. Staff redesigned the database to conform to the program business model. They added financial data and TANF eligibility data, and removed some data that was not needed. In hindsight, staff would have chosen a different query tool for the front end.

Data Sources

The data warehouse data source is the New England Child Support Enforcement System (NECSES), the child support automated system. NECSES also includes data from other sources, including TANF, the Department of Labor, the IV-E program (Foster Care), Maine Revenue Department, Financial Institute Data Match (FIDM), and Workers Compensation. The data warehouse includes some of these data in a summarized form.

The DSER User Group, which included BIS system analysts, DSER supervisors, and regional managers, determined the elements that would be included in the data warehouse. As DSER

makes enhancements to the data warehouse, other experts, such as enforcement agents, will be included.

Not all elements of the legacy system are in the data warehouse. For example, the enforcement subsystem elements are not included.

There is capacity to add new elements; since the data warehouse's inception, data have been added multiple times. Most recently, financial data were added. There are plans to add a multiple case indicator; receipt of collections, distributions, and disbursements; case change (e.g., from worker to worker or district to district); and case status change (e.g., from TANF to non-TANF). DSER also removed data that are not being used (e.g., financial and case data for cases closed longer than one and a half years).

Staff would like to include the enforcement diary (i.e., credit reporting, liens, garnishments, income withholding, notices of debt). This data is in NECSES and will be incorporated into the data warehouse in stages and as funds are available.

Data Uses

Federal Reporting	No
Performance Measurement	No
Self Assessments	Yes (sample only)
Ad hoc Capabilities	Yes
Types of Analyses	Point in time, trends Cycle time Projections

The primary function of a data warehouse is program management, including identification of problems. For example, the data warehouse is used to check data in NECSES to ensure its accuracy (e.g., data might indicate a non-custodial parent has two open employers—if so, cases are sent to the local offices for corrective actions). Staff can assess performance at the district and state levels. There are some pre-defined reports and queries. Many data warehouse users also can create ad hoc reports. Additionally, the data warehouse is used for self-assessment. Each is described below.

Pre-defined reports and queries. The data warehouse internet site—accessible to all DSER workers—includes a limited number of reports on performance and caseload management (e.g., cases closed, all children emancipated). Supervisors and regional managers use predefined reports to monitor staff performance (e.g., number of orders established, non-paying cases that should be transferred to another unit).

In addition to reports, data warehouse users have access to "canned queries." The queries are specific to an individual's caseload and can be used to monitor or assess performance. For example, one query identifies cases that meet license revocation criteria. Staff use the member ID number and the unique case ID number to track information across cases, individuals (a case can have multiple individuals), and/or orders.

Ad hoc reports. Managers and supervisors are able to query on any data field in the data warehouse. (Other DSER staff can request an ad hoc report from one of the DSER "super users.") They can select from a number of attributes, or data clusters. For example, a worker can examine his or her caseload for non-custodial parents that have not paid in 90 days, owe a

minimum of $5,000 in arrears, and have a confirmed employer. Or, he or she can identify cases that have not paid in a specified number of days (e.g., 30, 45, 90) and whether the non-custodial parent has an open employer.

Data are available at the individual, office, district, and state levels. Examples of questions commonly answered by the data warehouse include:

- Checking case status. Supervisors use the data warehouse to review workers' cases to determine if they are establishing a set amount of orders.
- Point in time comparison of case status, case change, address and employment histories.
- Cycle time analyses that indicate the amount of time that a case spends in a specific activity (e.g., locate). Dates are provided for comparison in history and change tables.
- Professional development. For example, if an office or worker is not closing cases, DSER may issue an instruction memorandum that training is needed.

Self-assessment. The data warehouse is used to produce a focused sample for all eight Federal self-assessment categories and to answer individual questions within several categories. Some questions cannot be answered because data are in "notepads" and must be read.

Policy Development. Staff use the data warehouse for research that may result in policy changes or modifications. For example, the state issues wage withholding orders for all cases. The data warehouse was used to demonstrate that staff could do income withholding for 99.98 percent of cases, but that for 0.02 percent there were locate and other issues. The state applied for a waiver for these cases because it demonstrated that performance would not improve by issuing withholding orders. In another example, Maine received another waiver from OCSE to include all cases in the IV-D system unless they actively opt out. An evaluation by the University of Southern Maine, using data warehouse data, found this policy was not cost-effective (e.g., it created cases that never came into the system). The policy was discontinued.

The data warehouse also has the ability to aggregate data to examine trends and patterns. For example, the data warehouse could be used to determine if a non-custodial parent moved every six months for the last two years, thus will likely move again in the near future.

The data warehouse is not used for Federal reporting. Audit data is drawn from NECSES. The data warehouse is not used to produce the OCSE 157, 34a or 396 reports. Local offices and individual workers cannot use the data warehouse to track their performance on incentive measures (i.e., there are no predefined reports or queries). However, supervisors and workers can request an ad hoc query.

Data Warehouse Users

There are different levels of data warehouse data access. All statewide DSER staff have access to predefined reports and canned queries. Currently, only supervisors and DSER management have unrestricted access to individual-level data and ad hoc capabilities. Data are accessed through a web-enabled interface.

Staff from other state agencies or other institutions (e.g., universities) cannot access the data warehouse. The University of Southern Maine's Muskee Institute used data warehouse-generated data for the waiver evaluation (described above), but received it in an ACCESS file.

Funding

The data warehouse was funded entirely by DSER, through the agency's general fund, incentive payments, and Federal financial fund participation.

The initial start-up costs were minimal and included the cost for software and limited BIS staff time.

The maintenance costs initially were about $1,100 per month for storage costs and $10.00 per user for license fees. With recent expansion, the storage cost increased to $5,600. DSER plans to reduce this amount by eliminating some closed case data.

Architecture

Design Issues	
How often are data collected/corrected?	Data loaded into production system every week; if there are problems, data are fixed in production system and refreshed on the weekend
How are data linked?	Data are linked using the Bi/Query models
How many tables are in final database design?	34
How much data are retained?	All data from select tables that are in the production system are kept in the data warehouse
What are the security requirements?	Access to individual-level data are limited. Currently only supervisors and DSER management have access.
Data Source Descriptions	
Hardware platform	Sun Fire V1280
Operating system	Sun Solaris Version 8.0
Database or file type	Oracle 9i
Extraction, Transformation and Loading	
Number of source systems	Two
Total data transported	30 gigabytes per month
End User Access	
Number of business areas in system	Two—Division of Support Enforcement Recovery, Bureau of Information Services
Tools used	Bi/Query and PL/SQL

MARYLAND CASE STUDY

Data Warehouse Snapshot

Date created	1998 (longitudinal system)
Primary uses	Policy development and program management
Primary data sources	Client Information System
Types of reports	Ad hoc, some Federal reports
Users	Family Welfare Research and Training Group staff
Primary funding source(s)	Federal Financial Participation (IV-D), TANF, University of Maryland Baltimore County

Background

The Family Welfare Research and Training Group (the Group) at the University of Maryland-Baltimore partners with the Maryland Department of Human Resources (DHR) to conduct welfare and child support research, policy development, and training. DHR houses both the child support enforcement and welfare agencies. For 20 years, DHR has funded a cross-program policy research initiative developed and operated by the Group. A data warehouse of sorts emerged from this initiative. The cross-program longitudinal database (the database) provides data that helps DHR program managers and front-line staff implement and evaluate welfare and child support programs.

The Group involved DHR in the database plans because the agency owns and houses the majority of the data. Meetings were informal and provided an opportunity to discuss why the data was needed. Through participation in policy meetings and trainings, the Group understands the issues that are of greatest concern to DHR staff. DHR continues to support the database because staff see the value of data analysis; additionally, the resulting reports "make the agency look good in front of the governor and assembly."

Implementation

As noted above, the database was implemented over many years. The initiative began in 1982 and used paper-based administrative data. In 1995, DHR began sending month-end welfare and Food Stamp files disk. These data indicated which cases were active and which ones were closed. In 1998 DHR started sending quarterly child support data, in addition to Medicaid, Food Stamp and TANF data. This was the genesis of the longitudinal database.

Data Sources

The primary data source for the database is the Client Information System (CIS), which is the umbrella system for the TANF, child support, Medicaid, and Food Stamp programs. CIS data include common identifiers (e.g., name, Social Security Number, Date of Birth, and IRN—a unique number) that can be used to link clients across programs. The data also include child support and TANF case action logs. DHR sends the CIS data monthly.

The database also includes data from:

- Maryland Directory of New Hires (received monthly, linked by Social Security Number).

- Child Care program data from DHR (received monthly, linked by SSN).
- Work Opportunities (Maryland's Welfare-to-Work program) data from DHR (received monthly, linked by SSN).
- Hard copies of all voluntary in-hospital paternity acknowledgements since 1996 (from the Division of Vital Records). Data are entered into the database then linked by mother's or father's SSN.

In addition to data received directly from the state government, the Group has an agreement with the University of Baltimore (the state's archivist for Unemployment Insurance data) to get quarterly wage data for a sample of SSNs.[8]

Data are available at the individual, office, county, and state levels.

Additional data can be added to the database. Child support and public safety (corrections) officials and staff from the Group are discussing creation of a 10-year file on incarcerated obligors. Public safety would provide child support with a monthly updated file.

Data Uses

Federal Reporting	Yes
Performance Measurement	No
Self Assessments	No
Ad hoc Capabilities	Yes
Types of Analyses	Point in Time Point-in-Time Comparisons Longitudinal

The database is used for policy development and program management. Most reports are ad hoc in nature, although the database is used to create some Federal reports for the IV-D agency.

Federal reports. The database is used to produce the OCSE 34a and 157 reports. The Group currently runs the data, but is in the process of transitioning the task to the IV-D agency. The Group will train the relevant child support staff.

Other reports. In addition, the Group creates monthly statistics reports on topics such as the Driver's License Suspension Program and the Paternity Affidavit Program.

Ad hoc reports. Most of the reports produced are ad hoc in nature and respond to specific needs of IV-D and IV-A officials who use the data for program management.

Between August 2003 and April 2004, there were 499 ad hoc report requests from child support officials. Local IV-D program managers or lead supervisors at the central child support office can request ad hoc reports (they can email or fax a request form to the Group). Generally, the reports answer case management-oriented questions (e.g., how many cases in a given office are in the locate function for a specified amount of time). Reports do not track local performance on incentive measures, but a local office may request a report on the number of cases not paying, the number of cases in arrears, or other measures related to the incentives. As appropriate, managers will give their staff case information from the reports for follow-up work.

[8] The Group indicates which SSNs are needed for the study.

The Group also conducts guidelines reviews. In addition, IV-D directors from the state's "demonstration counties" (the larger Maryland counties) met with the Group to discuss creating regular reports that will help them better manage their programs. These county-specific reports would contain a variety of information concerning the caseload. Examples include quarterly reports on cases without orders and cases with orders but not payments in the past 90 days.

Policy research. The IV-D director is interested in using the database for policy-oriented research. For example, a study of arrears among incarcerated obligors is in the planning stages. A 2002 study examined how many children born out-of-wedlock for whom a paternity acknowledgement was filed became known to the child support program within one year, and how many became known to the TANF, Medicaid, or Food Stamp programs during the same period. (The data analysis found that one-third of children whose paternity is acknowledged become known to the child support program; 75 percent become known to other DHR programs.) This study informed cross-program planning between child support, TANF, and other programs that serve low-income families.

An earlier study conducted found that about 80 percent of welfare recipients in Baltimore failed to show up for intake interviews with child support staff. Although a large proportion of the welfare recipients received a letter, they did not know what it meant and where to go; additionally, there was no consequence for failing to attend the intake meeting. As a result of these findings, the IV-D office piloted a program in which a IV-D worker was outstationed in a TANF office. Subsequent research showed a sharp contrast between the statewide no-show rate (75%) and the rate in the Baltimore office (5%).

The database's value extends to other program areas as well, such as the largest longitudinal study of welfare leavers in the country. The study began in October 1996 (the first month in which welfare reform in Maryland took effect). Every month staff draw a five percent sample of closed welfare cases and track the individuals and cases moving forward. The results have appeared in a series of eight *Life After Welfare* reports. The most recent, published in October 2003, provides information about the demographic characteristics, welfare receipt patterns, and post-exit employment patterns of the 8,567 welfare leavers and their families researchers have been tracking since the study began.[9]

Database Users

Access to the database is highly restricted. There are four programmers from the Group who have access to all of the data (i.e., data with identifiers). Three study directors who design methodology and write data specifications for the lead programmer have access to study-specific data, including identifying information. Research analysts at the Group work with data stripped of identifiers.

Data are accessed through a dedicated server. The lead programmer creates a folder for each study and permission is limited to the people working on the study.

Data are not accessible to other state agencies or university staff beyond those in the Group.

[9] Other studies include *Life Without Welfare: The Prevalence and Outcomes of Diversion Strategies in Maryland; Caseload Exit series* (characteristics of cases closing during the study year); *Life on Welfare series* (a snapshot of active cases, changes in the caseload over time); *Setting the Baseline series* (child welfare entries among welfare exiters, patterns of welfare recidivism).

Funding

The Group has two interagency agreements with the child support enforcement program and the Family Investment Administration (TANF)—both within DHR. Very little funding, however, comes from the state (e.g., the General Fund). Most of the Group's funding comes from the TANF block grant and the IV-D Federal Financial Participation (FFP) funds, with the University of Maryland-Baltimore County contributing the non-Federal share. The Group also seeks Federal grants. One example is a three-year grant from the Federal Office of Children, Youth and Families to study client assessment as performed by local welfare caseworkers. Another example is a grant received from the U.S. Department of Health and Human Services' Office of the Assistant Secretary for Planning and Evaluation to study the characteristics and circumstances of the current TANF caseload.

Because the database evolved over time, staff did not estimate the cost of building it. However, staff estimate that the cost of the ongoing child support research is about $600,000 to $700,000, while the cost of the 20 ongoing TANF-related studies is about $1 million.

Architecture

Design Issues	
How often are data collected/corrected?	Child Support, Work Opportunities, Food Stamp and TANF Data come monthly, UI data arrives quarterly.
How are data linked?	Data are linked by SSN, IRN (a unique identifier), or Name and date of birth, depending on the source of the data
How many tables are in final database design?	514
How much data are retained?	All data from the source systems are retained
What are the security requirements?	Access to the database is highly restricted. Four programmers and three study directors have access to individual-level data; analysts work with de-identified data. Others submit requests for data.
Data Source Descriptions	
Hardware platform	Intel x86
Operating system	Windows NT 2000
Database or file type	MS SQL Server
Extraction, Transformation and Loading	
Number of source systems	Four
Total data transported	194 gigabytes per month
End User Access	
Number of business areas in system	Just the researchers. All requests for information go through them.
Tools used	One, direct SQL queries into the system from one of four databases

MICHIGAN CASE STUDY

Data Warehouse Snapshot

Date created	1998 (statewide)
Primary uses	Parent locate, Federal reporting
Primary data sources	Child support automated system (MiCSES)
Types of reports	Federal, ad hoc
Users	All IV-D staff
Primary funding source(s)	Federal FFP

Background

Michigan's data warehouse was deployed statewide in September 1998. The impetus for creating it was the need for a state-wide computer system to satisfy provisions of the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA), related to child support enforcement program automation.[10] The state previously had a distributed database design in which local county data were stored in separate databases. It was difficult to use the data for Federal reporting. The data warehouse provides Michigan with a single repository for child support data and was certified as the state-wide system. In addition, the data warehouse integrates child support data with key elements from other state agencies and external sources (e.g., financial institutions), thereby enhancing parent locate capabilities.

The primary staff involved in the decision to create the data warehouse were from the state Office of Child Support (OCS) and Child Support Systems (overseer of the state data systems for the past 19 years).

Implementation

Planning for the data warehouse started in March 1997. Michigan launched the prototype in Muskegon County in August 1998.[11] It was deployed statewide in September 1998.

Legislative requirements (i.e., the 1988 Family Support Act, PRWORA, and state laws that authorized child support to access large amounts of data) ensured that OCS had relatively little difficulty in securing cooperation from other agencies. The exception was the Secretary of State's Department of Motor Vehicles (DMV), which was reluctant to collect Social Security Numbers (SSNs) and link them to driver's license numbers. Previously, DMV collected SSNs only for commercial vehicles. The Secretary of State appealed the requirement to Federal Court and lost. As of May 7, 2004, the DMV collects SSNs.

According to data warehouse staff, the schedule for creating and implementing the data warehouse was reasonable and workable and there were no unexpected obstacles in implementing the design.

---

[10] Under PRWORA, states must have in effect a statewide automated data processing and information retrieval system, which by October 1, 1997, meets all the requirements of Title IV-D of the Social Security Act enacted on or before the date of enactment of the Family Support Act of 1988, and by October 1, 2000, meets all the title IV-D requirements enacted under PRWORA.

[11] Muskegon was selected because of existing contacts with the state office and because they were willing to participate.

In hindsight, staff would have changed the current design of the data warehouse to make it more useful by better capturing the entire child support database. The current data warehouse captures only the crucial child support data needed for the Federal Case Registry interface, the Federal reports (e.g., the OCSE 157) and non-custodial parent location. This would have allowed much greater flexibility for enhanced functionality.

Data Sources

The data warehouse includes both state and Federal data from a variety of agencies.

The state data sources are:

- MiCSES, the statewide child support enforcement system, is used for enforcement, locate, and Federal reporting.
- Department of Treasury: The Michigan employer file includes Federal tax IDs, employer names, employer addresses, and is used as the basis of employer information.
- Department of Corrections: data is used for locate and for modifications.
- Department of Natural Resources (hunting, fishing, off road vehicle licenses): data are used for locate and lien purposes.
- Workers Compensation (contested only) claim history: data, including name and address data, are used for locate purposes.
- State Directory of New Hires is used for locate purposes and information is also used as basis for the Income Withholding Notice (IWN).
- State Police: Personal Protection Orders (PPO) filed in Michigan are used to set the family violence indicator. This information is highly restricted and cannot be used for other purposes (e.g., locate).
- Financial Institution Data Match (FIDM): financial account information for delinquent obligors is matched to the Michigan caseload against Michigan institutions and is used for lien activity.
- Department of State (Secretary of State): Driver's license history is used for locate, license suspension.
- Bureau of Workers and Unemployment Compensation: Unemployment data –including name and address data—are used for locate and IWN purposes.
- Department of Community Health: includes Medicaid and is used for Medicaid commercial recovery.
- Department of Consumer and Industry Services (CIS): Commercial license data—including professional licenses issued by Michigan—are used for enforcement and locate.
- MiCSES Medical Interface: Health insurance coverage information (general medical, dental, vision) is received from nine insurers and used for medical support enforcement and Medicaid commercial recovery.

The Federal data sources are:

- Federal Parent Locator Service (FPLS): the data warehouse acts as the intermediary in FPLS locate requests between OCS and the Federal government. FPLS data passing through the data warehouse are retained for locate activity.
- Federal Case Registry (FCR) / Social Security Administration (SSA) (for SSN corroboration, it is sent back through the FCR): The state caseload is submitted to the FCR and proactive matches (other state child support cases) are returned to the data warehouse. The data are used for locate and interstate purposes, and name, date of birth, and SSN verification (via SSA). The information is also used as the basis for Federal reporting (OCSE 34a and 157).
- National Directory of New Hires (NDNH): new hire, quarterly wage, and unemployment data based on submission to FCR is used for locate and IWN.
- Multi-State Financial Institution Data Match: financial account information for delinquent obligors matched to Michigan caseload (against multi-state institutions) is used for lien activity.

Data are available at the individual, case and county levels.

Data warehouse data elements were selected based on their ability to support specific functions: parent locate, child support enforcement, Federal reporting, FPLS, and family violence determination. The data warehouse does not include detailed financial information or enforcement information (e.g., court dates, bench warrants, income withholding notices). There are also no transactional data.

The data warehouse has the capacity to add new elements as authorized by OCS management. The long term plan includes a formal requirements analysis. Some of the additional or improved sources OCS anticipates adding include:

- Michigan Department of Treasury: non-financial tax information from the 1040 (SSN, name, and address).
- Department of Natural Resources: restart campground reservations (for lien purposes).
- Bureau of Workers and Unemployment Compensation: claim history for uncontested claims.
- Department of Environmental Quality: mineral rights information (name, address, and wage records of anyone who is paid for mining).

Data Uses

Federal Reporting	Yes
Performance Measurement	Yes
Self Assessments	No
Ad hoc Capabilities	Yes
Types of Analyses	None at this time

The primary functions of the data warehouse are parent locate and Federal reporting. The data warehouse also interfaces with the FPLS and FCR and serves as the state directory of new hires and the source of the data feed to the National Directory of New Hires. The data warehouse is used to locate information on parents' addresses, employers, and quarterly wages. Using a date

or birth, SSN, driver's license number, or other identifier (e.g., name) staff can use the data warehouse to search the multiple data sources for address and employer information. A "hit" for an address and/or employer will indicate the source of the information (e.g., FCR, state directory of new hires), the effective date (e.g., when employment began), the field matched on (e.g., SSN), and the source date (e.g., how old the data are). The case worker can download the information into MiCSES and include written notes in the case history. The location status will be changed in MiCSES. The system automatically generates a letter to the employer to verify employment and/or a notice to the post office to verify the address.

Staff can also search data with date ranges. For example, staff can look at a parent's employer history, sorted by most recent source date. Staff can determine where the parent was working, for how long, and how much he or she earned in a quarter.

Federal reporting. The data warehouse produces the OCSE 157 and 34a quarterly. The data warehouse is used to produce the Data Reliability Audit detail (i.e., the data that Federal auditors use to audit the OCSE report forms).

County performance. State-level staff use the data from the OCSE 157 report to track local office performance. The state distributes incentives to individual counties based on their performance. In addition, the data warehouse allows local offices to track their own performance in each of the five Federal measures.

OCS uses data from the OCSE 157 and 34a, and the IV-D caseload report to administer CRP (Cooperative Reimbursement Program) contracts, which pay for local Friend of the Court (FOC) offices. FOC administers the child support program at the county level.

Pre-defined queries. Data warehouse users can access a number of predefined queries, including:

- Person locate: People and address; employment and income; assets and licenses.
- FOC-specific queries: Tax offset; data range queries for caseworkers; arrearage range; court order.

Ad hoc reports. The data warehouse has ad hoc capabilities. Types of ad hoc reports that are requested include:

- List of non-IV-D cases.
- List of children born out-of-wedlock without paternity established.
- Count and summation of payers in arrears that are incarcerated or deceased.
- Payers in arrears that have/have not paid in the last six months.

Targeted ad hoc enforcement reports are available to support new or innovative practices. For example, a county might want to try a new method to increase collections, such as suspending deer hunting licenses in September—the beginning of hunting season.

Any authorized user can request an ad hoc report through the help desk. OCS approves and prioritizes the requests. Sometimes OCS management will request a report (e.g., the current national debt study that explores debt compared to arrearage levels).

Data Warehouse Users

All child support staff and technical staff have access to the data warehouse (3,276 users). Only OCS technical staff have unrestricted query access. Non-technical staff (e.g., caseworkers, managers) have read-only access. Different users have different levels of viewing access. For example, the Attorney General's office, which has a contract with OCS to handle criminal non-support and contested cases, has an unrestricted view of the data. Others have more limited access. FOC case workers can only access the screens that are related to their jobs. Prosecuting attorneys and support specialists have the same view as FOC case workers. OCS provides training through formal classes, peer-user training, and self-directed courses. Users can contact the help desk for assistance with ad hoc requests.

Some non-IV-D state staff have access to the data (e.g., the Department of Community Health has a limited view of health data, the Office of the Inspector General and the Bureau of Workers and Unemployment Compensation have access for fraud detection purposes). Users access data through a client / server application. Access is granted as appropriate to the job.

Funding

The data warehouse is supported by Federal funds (Federal Financial Participation, or FFP) as part of the statewide child support system.

Architecture

Design Issues	
How often are data collected/corrected?	Daily, weekly, monthly feeds are received; data cleaning, especially names/addresses integral part of collection process.
How are data linked?	Data are linked in the data warehouse through extensive matching
How many tables are in final database design?	465
How much data are retained?	All data since 1998; closed and terminated cases are retained; data is neither purged nor deleted.
What are the security requirements?	Users sign a propose use and non-disclosure agreement; two log-ins are required and access is only allowed from within the State of Michigan network. Data access is restricted by role via views. Only technical staff have unrestricted query access (with identifiers); others only have viewing access.
Data Source Descriptions	
Hardware platform	NCR WorldMark 5380
Operating system	NCR MP/RAS (UNIX)
Database or file type	Teradata V2R5
Extraction, Transformation and Loading	
Number of source systems	14
Total data transported	7.5 gigabytes per month; entire data warehouse is 319.1 gigabytes
End User Access	
Number of business areas in system	Eight: Friend of the Court, Office of Child Support, prosecuting attorneys, child support specialists, Attorney General, Department of Community Health, Office of Inspector General, Unemployment Agency. Each has a unique view into the data.
Tools used	BI-Query applications, Queryman, child support source system application

NORTH CAROLINA CASE STUDY

Data Warehouse Snapshot

Date created	1999
Primary uses	Federal reporting, program management
Primary data sources	Child support, TANF, Medicaid, Food Stamp systems
Types of reports	Federal reports, ad hoc reports
Users	Staff from the Divisions of Social Services, Mental Health, Child Development, Public Health, and Aging
Primary funding source(s)	TANF

Background

The North Carolina Department of Health and Human Services' (DHHS) data warehouse was created in 1999. The impetus was TANF reporting requirements, and it was funded with TANF dollars. However, the data warehouse now contains data from and serves a number of programs, including child support. The data warehouse can track participants across programs.

The DHHS Division of Social Services (DSS) policy staff, business personnel, and TANF administrators at the state and county levels were involved in the decision to create the data warehouse. There were no serious obstacles to overcome as staff believed that the data warehouse would make Federal reporting easier. The fact that staff would receive training on how to use the data warehouse was also a selling point.

Implementation

DSS is rolling out the data warehouse in phases. There were nine months between the beginning of the planning process and the rollout. In general, the schedule was reasonable and the agency did not encounter any major obstacles.

In hindsight, DSS staff indicated that they would have added additional capabilities to summarize and aggregate detail data. They would have also liked to see greater focus on customer tracking across different programs' systems (e.g., TANF, child support, Food Stamps).

Data Sources

The data warehouse serves a number of DHHS programs. Data are supplied by five "divisions":

- Ninety percent of the data are from DSS; child support accounts for about 50 percent of this data. The remainder of DSS data is from the TANF, Medicaid, and Food Stamps programs; services (e.g., transportation); central registry (child welfare); and employment services.
- The remaining 10 percent of the data is from four divisions: Division of Mental Health (mental health treatment, clinical records), Division of Child Development (Head Start, child care, regulatory entities that monitor child care providers), Division of Public Health (AIDS treatment), and the Division of Aging (guardianship program).

Each division has its own source system(s). The data warehouse is independent of these systems. However, all of the programs' physical machines are under one roof (the data center)

with the exception of the Division of Aging's source system. The DHHS technical staff extract the data from each system and load it into the data warehouse on a cyclical basis—daily, monthly, or quarterly, depending on the source. The extraction, transformation, and loading of data are an automated process: The "trigger file" indicates when the source has completed its "closeout" process and tells the data warehouse when data are ready to be added. Data are available at the employee, client, local office, county, and state levels.

Potential users and technical staff had input into the elements included in the data warehouse. Field, central office, and technical staff from DSS examined the source systems, and discussed what types of reports could be generated with the data warehouse given the available data and the programs' reporting needs (e.g., the OCSE 157 report and self-assessment reports for the child support program staff).

New elements can be added to the data warehouse. A steering committee meets bi-weekly and continues to identify data that programs want to include. Also, local offices make recommendations to the committee. Hardware and time constrain data additions. In addition to the programs already included in the data warehouse, some of the smaller DSS programs (e.g., energy assistance) have expressed an interest in including data.

Child support has not yet moved all of its data into the data warehouse. Currently, the data warehouse contains the data necessary to program the OCSE 157 report and all of the self-assessment reports. The ultimate goal is for the data warehouse to produce all OCSE reports. The data warehouse contains case, enforcement, and financial information, and a project is in progress to add income withholding data. The Legacy mainframe will be used only for manual reviews of data for testing purposes.

Data Uses

Federal Reporting	Yes, 157
Performance Measurement	Yes
Self Assessments	Yes
Ad hoc Capabilities	Yes
Types of Analyses	Point in time, trends, cycle time

The primary function of the data warehouse is the Federal reporting requirements for the five agencies' programs: DSS (TANF, child support, Food Stamps, employment programs, Medicaid, child welfare), mental health, child development (e.g., head start), public health (AIDS treatment), and aging.

Predefined reports. The data are used by program staff to monitor and manage their respective programs. The data warehouse includes 560 reports across the different programs. These reports include canned queries as well as queries created by individual users. About 60 of the canned queries are child support-related.

Ad-hoc reports. The Child Support Standard Reports category contains queries developed by users. Users can query any data element. If the creators of these queries thought other users could benefit from them, they publishe them and make them available to all users. For these queries, the user specifies the parameters (e.g., month). In most cases, there is up to five years of data. Users can run their own queries every month when the data are updated. These queries are used by program staff for case clean-up, performance tracking, and training needs.

These customized queries are often created by local offices and are county-specific. For example:

- Counties can execute a query to show the total count of court actions initiated in a specific time frame and show the total count of those actions served/unserved to validate their need for a deputy to be assigned to child support.
- A county might need demographic information for a language study regarding the demographic growth of their county.
- A local plant might close and a county executes a query to see how many cases presently have an income withholding with this employer to see the forthcoming impact to their collections.
- A county with a military base might query how many interstate cases they have with a certain country overseas.

County program directors use the data for "friendly competition." They contend that worker productivity improves because their performance is measured and compared to that of other workers. Regional consultants and trainers use the data warehouse to track local office performance to determine if trainings are needed. The data warehouse also helps local managers determine if staff need to be moved from one function to another (e.g., enforcement to establishment). Data are also available at the individual worker level, allowing managers to monitor the performance of individual staff members. Local offices and individual workers can track their performance on incentive measures (except cost-effectiveness).

In addition, child support staff use the data warehouse for the following reports:

Federal reports. The data warehouse produces the OCSE 157 report on a monthly basis.

Self-assessment reports. The eight self-assessment reports are updated quarterly.

Data Warehouse Users

Overall, 2,800 people have access to the data warehouse. This includes staff in DSS across the five divisions, partner organizations (e.g., Food Stamp fraud detection), and 10 outside organizations that support DSS.[12] About 250 of the users are in the child support program.

Data are accessed through a web-enabled interface. Access to the data warehouse is organized along program/organization lines. DSS has security officers who determine whether or not a given individual should be granted access to specific data within the data warehouse. Generally, staff only have access to data related to their particular program. Within the user group (e.g., child support), users can access canned queries and write their own reports. Child support data users have access to individual-level data that includes Social Security Numbers. Each individual program was responsible for training its staff. DHHS also operates a central help desk all users can go to for assistance with reports and writing queries.

Some organizations outside of DHHS have access to the data. As noted above, there are 10 outside organizations that support various DSS functions. For example, the Jordan Institute at

[12] For example, Medicaid has a contract with the Professional Consulting Group, a contractor that assists with medical support enforcement by comparing employers of non-custodial parents against a database of health insurance companies to determine if health insurance is available.

the University of North Carolina has access to the TANF data because staff conduct policy studies for DSS and collect data used in TANF reports. No other state agencies currently have access. Legislative actions, however, might necessitate broader access; the legislature is considering a study of how children with mental retardation fare in the education system.

Funding

The primary funding source was TANF.

Architecture

Design Issues	
How often are data collected/corrected?	The feeds depend on which database the information comes from. Data are corrected as mistakes are noticed and reported. Child Support data are loaded monthly. Food Stamps (EBT) is loaded daily. Most of the TANF related is loaded monthly
How are data linked?	All of the different tables are linked by an ID assigned to the client. For mental health data, the tables are linked by the client's name. Typically tables are linked by either case or individual. Linkage is system specific.
How many tables are in final database design?	240
How much data are retained?	The data warehouse keeps data for 5 years. Due to hardware limitations, five years of case history is retained on the live systems, the rest is archived.
What are the security requirements?	Access is controlled through the web via Business Objects, and permissions are assigned to users based on their needs. Data are restricted by role, or organizational boundaries.
Data Source Descriptions	
Hardware platform	Web servers: Compaq x86 Database: Sun E10K
Operating system	Windows NT4 and 2000 Server and Solaris
Database or file type	Sybase ASIQ
Extraction, Transformation and Loading	
Number of source systems	Four
Total data transported	150 gigabytes per month
End User Access	
Number of business areas in system	• Social Services (Child Support, Food Stamps. Child Welfare) • Mental Health • Aging • Public Health • Outside Agencies: US FDA, UA OIG, the Professional Consulting Group and others.
Tools used	Business Objects via Web Intelligence

PENNSYLVANIA CASE STUDY

Data Warehouse Snapshot

Date created	1998
Primary uses	Federal reporting, program management
Primary data sources	Child support automated system (PACSES)
Types of reports	Federal reports
Users	IV-D staff
Primary funding source(s)	Federal (IV-D Incentive Funds), State (Division of Automated Child Support Enforcement Systems)

Background

Pennsylvania's child support data warehouse was created in 1999 to expand reporting and querying capabilities for the Pennsylvania Child Support Enforcement System (PACSES), the state's automated child support system, and to support the creation of Federal reports.

The child support warehouse resides on the same server as databases for other Department of Public Welfare (DPW) programs, such as TANF. The data are not currently linked across programs.

Implementation

The child support data warehouse was implemented incrementally, following the standard data warehouse system development lifecycle for each individual reporting structure. As such, the system grew report-by-report as functional and technical resources became available.

With the exception of the Federal reports (OCSE 34a and 157), which had fixed deadlines based on the Federal reporting cycles, PASCES used this "time and materials" approach in implementing reports, taking as much time as was required to ensure the reporting logic was validated. This process was somewhat slowed due to the limited availability of state staff to identify requirements and technical resources to implement the reports.

Another challenge in designing and implementing the data warehouse was ensuring compatibility with PACSES's mainframe system while still satisfying the more general Federal reporting requirements. Compatibility issues emerged when staff developed processes to ensure that any changes to the state system would be reflected in the data warehouse.

The implementation process also required developing the system while standards (e.g., data naming standards) and technology platform (tools, software versions) were still being established. In hindsight, staff would have liked to create and vet comprehensive standards and technology platforms before beginning development.

Data Sources

The data warehouse data originate in the child support mainframe, PACSES. Data are transmitted to the data warehouse monthly and kept for three years.

PACSES (but not the data warehouse) interfaces with TANF, Unemployment Compensation, Department of Revenue (state tax and Federal taxes), the Department of Motor Vehicles, the credit bureau, the State and Federal New Hire Directories, the state disbursement unit, the professional license bureau, the Federal Case Registry, and the Federal Offset Program. Data is linked in PACSES using client's Social Security Number, date of birth, name, or a combination of these identifiers.

To determine which elements to include in the data warehouse, PACSES staff modeled the business case for the primary data warehouse functions: Federal reporting and self assessment. The data support these functions. About 95 percent of elements from the legacy system are included in the data warehouse. Data are available at the state, county, case, and member level.

There is capacity to add new elements to the data warehouse. The Data Management Work Group, which includes representatives from PACSES, six counties, and state Bureau of Child Support Enforcement (BCSE) staff, meets periodically to discuss additions and enhancements. They are currently discussing whether to add information from hospitals on paternity establishment.

Data Uses

Federal Reporting	Yes
Performance Measurement	Yes
Self Assessments	Yes
Ad hoc Capabilities	Yes
Types of Analyses	Point in time, trends

The primary function of the data warehouse is to create Federal reports, conduct annual self-assessments, and assess program performance in specific subject areas (financials, enforcement, establishment). BCSE and county IV-D programs use the data warehouse to track four of the five Federal performance measures (paternity establishment, cases with orders, current collections, cases paying toward past-due support).

In addition, the data warehouse supports data integrity. When data are loaded into the system, staff can identify problem cases. In one instance, 40,000 records showed children who had been born in the future (there was a problem with the century indicator conversion). BCSE used the data warehouse to find irrelevant or illogical cases, which were sent to the appropriate counties for clean-up work.

In the future, BCSE plans to use the data warehouse more proactively to identify problems and ask counties to respond. For example, BCSE has Subject Matter Experts (e.g., specialists in paternity establishment, order establishment, enforcement, review and adjustment), each of whom is responsible for working with 10 counties. Ideally, these staff would monitor their assigned counties' performance and contact county directors to learn about best practices (if the county is doing well) or provide technical assistance (if the county is underperforming).

Federal reports. The data warehouse is used to generate the following Federal reports:

- The OCSE 157 (which is divided into seven reports; Section A, case inventory; Section B, paternity establishment; Section C, services requested; Section D, services provided; Section E, medical support; Section F, collections due and dispersed; Section I, non cooperation). The user can examine the data behind the reports. For example, the

Section A Report provides a general inventory of the case and support order load. The user can view each line item (e.g., Line 1, cases open at the end of the fiscal year) by county, Federal fiscal year, and case type (e.g., TANF, non-TANF, Foster Care).

- The OCSE 34a (quarterly Federal report, monthly collection report by county, monthly disbursements report by county). The user can examine the data behind the reports (contained in "cubes"). For example, the user can click on a link to the "collection cube" and view collection amounts for IV-D and non-IV-D cases by payment source (e.g., IRS Tax Refund Offset), geography (e.g., county), and case category.

Self-assessment reports. The data warehouse creates monthly reports on the eight self-assessment dimensions (establishment of paternity and support orders, enforcement of orders, disbursement of collections, review and adjustment of cases, case closures, establishment and enforcement of medical support orders, interstate services, expedited process). Reports can be sorted by geography (e.g., region, county), order type, and case category. A number of the self-assessment reports provide "drill-through" capabilities where a user can filter the data using various dimensions and then generate an "actionable" case list, which is then provided to a worker for further analysis/action.

Enforcement remedies reports. The data warehouse creates the following enforcement-related reports:

- Drivers license suspension report. This report sums the enforcement activity and payments attributable to the driver's license suspension program. The report shows payment amount across a number of dimensions: geography (county, region, class[13]), time (Federal fiscal year, state fiscal year, current month), enforcement action (notice of intent to suspend, order to suspend, opportunity to contest, contest, order to reinstate), case type (disability/SSI, foster care, general assistance, medical need only, TANF, non-TANF), case category (child support, educational support), case status (archived, closed, open), interstate, and establishment type (paternity, support order, order established). An example: As of July 2003, how much money was collected by way of wage attachment from Class 3 counties after driver's license suspension?

- New Hires collection report. This report details collections directly attributable to state and national new hires programs (e.g., collections from Pennsylvania New Hire, Quarterly Wage Report, Unemployment Compensation, and W-4 Records). Data can be viewed across the same dimensions as the driver's license suspension report.

- Financial Institution Data Match (FIDM) report. This report indicates the amount collected in-state and multi-state, the number of FIDM collection receipts, and the average FIDM collection. Data can be viewed across the following dimensions: geography, time, case type, FIDM type.

- Credit Bureau report. This report indicates the number of non-custodial parents who have not met their child support obligations and are eligible for credit reporting. Data can be viewed across the following dimensions: geography, time, case category, case type, case status, interstate.

[13] Class indicates the size of the caseload. For example, Philadelphia County would be Class 1; Allegheny County (Pittsburgh) would be Class 2.

Other reports. Data warehouse staff are creating a County Annual Report that will provide county managers with key information to help them manage their operations and measure their performance against established standards and other counties' performance. The report will contain information about performance on four incentive measures,[14] financials (e.g., undistributed collections, distributed collections, disbursements by payment source), cases (case type, status, activity), and bench warrants. It is currently being tested by the Data Management Work Group.

A monthly report card for management-level staff also summarizes collections by month and method (e.g., wage attachment, drivers' license suspension) and history (i.e., where the state was at the same time last year).

Cubes. Cubes allow data warehouse users to sort and filter data to attain the desired information. Cubes can be accessed from the main screen or through the relevant Federal or enforcement report. The subject areas of the cubes are: attachments, cases, collections, disbursements, distributions, member roles, members, nondistributed collections, obligations, and support orders. Information can be filtered across a number of dimensions (e.g., geography, date, case category, TANF status, establishment type). Examples: As of June 2003, how many cases were submitted to the IRS for tax return intercept in Allegheny County (Cases cube)? As of July 2003, how many defendants in Philadelphia were ordered to provide medical insurance but were unable to do so (Support Orders cube)?

Ad hoc capabilities. County staff and BCSE staff can request ad hoc reports by submitting to PACSES a Data Processing Service Request. For example, a staff person might request a report on all the cases in the specified county where the non-custodial parent is dead and there are no arrears.

Data Warehouse Users

About 350 users have access to the data warehouse. These include management-level staff in the 67 counties, as well as BCSE and PACSES staff. Generally, BCSE staff use the data warehouse for Federal reporting, while county staff use it for analysis and reporting on business operations. Access to data is more limited for county staff.

Users access the data warehouse through a password-protected web-enabled interface. Once they are in the system, the main screen directs users to reports or cubes.

At this time, non-IV-D staff do not have access to the child support data in the data warehouse. Data are not currently linked across program databases. There are several issues preventing such a linkage—siloed program offices that each have separate funding, priorities, and business requirements. In addition, evolving standards have resulted in basic design discrepancies which would make it both costly and complex to integrate several program offices in a single data store.

Funding

The data warehouse was funded by both state and Federal funds. There was no specific allocation of funds by the state for its creation. Staff estimate the data warehouse cost about $10 million to build. Infrastructure and maintenance costs are shared across program offices, and labor costs are associated with staff dedicated to each program office.

[14] Paternity establishment, cases with orders, collection on current support due, cases paying toward past-due support.

Architecture

Design Issues	
How often are data collected/corrected?	Collected monthly; corrected only if necessary for Federal reports and the annual data reliability audit.
How are data linked?	Linked in the child support source system (PACES) prior to data warehouse transfer. Data are linked by SSN, Date of Birth and/or client name
How many tables are in final database design?	64
How much data are retained?	Data warehouse retains data for three full years
What are the security requirements?	Full hardware firewall and no access to the database from the outside. Additionally, COGNOS security is used to limit access to needed views, data and reports.
Data Source Descriptions	
Hardware platform	Sun Fire 15K for the database and Unisys ES7000 for the webserver
Operating system	Solaris/Windows 2000
Database or file type	Oracle 9i
Extraction, Transformation and Loading	
Number of source systems	One
Total data transported	11 to 12 gigabytes per month
End User Access	
Number of business areas in system	Three: Bureau of Child Support Enforcement, PACES, county-level staff
Tools used	COGNOS, ESRI

VERMONT CASE STUDY

Data Warehouse Snapshot

Date created	2002 (initial work began)
Primary uses	Program management Policy development
Primary data sources	Child support automated system (ACCESS)
Types of reports	(to be added)
Users	Office of Child Support staff
Primary funding source(s)	OCSE 1115 grant

Background

Vermont's Decision Support System (DSS) is called PEAKS (Performance Enhancement and Knowledge System). The requirements analysis began in September 2002 and identified the informational, analytical, and functional needs that were candidates for inclusion in the DSS.

The Office of Child Support's (OCS) impetus for creating PEAKS was the need to access integrated, timely, and accurate data from multiple sources in order to enhance program performance and management at both the state- and individual-worker level. According to Jeffrey Cohen, the IV-D director, the goal of the data warehouse is "to increase productivity and effectiveness of all state staff by allowing users to convert data into information that is accessible and understandable, resulting in actionable activity." Also, the data warehouse allows OCS to leverage the wealth of information currently contained within the legacy transaction system.

Two key state-level staff were involved in the decision to create the data warehouse: Jeffrey Cohen and Cindy Griffith, the Vermont Administrator and DSS Project Manager. Potential IV-D users offered input. IV-D staff needed to work with their counterparts in the IV-A office because the latter manages and maintains the ACCESS mainframe system, which includes child support data.

Implementation

The state received an 1115 demonstration grant from OCSE to build a data warehouse over a 23 month period. Work concluded in September 2004. OCS's time line incorporated five discrete iterations, each of which adds data and/or functioning capabilities:

1) Establish a foundation and framework for the DSS. Case data and the web-enabled on-line analytical processing (OLAP) functional capability are added. Staff will be able to manipulate and analyze data on line.
2) No new data are added, but new functionalities include data mining capabilities and pre-defined reports.
3) Financial data are added. New functionalities include the "digital dashboard" (a summary of performance information) and data mining capabilities.
4) and 5) Additional data and data mining capabilities are added.

The first iteration was completed in April 2003, the second in June 2003, the third in October 2003, the fourth in March 2004, and the fifth in September 2004.

In hindsight, the state would have added more training for staff, including training for system staff remaining on site after the contractor's work is complete. Staff also suggested that they would have liked to implement the data warehouse in six iterations. Between iteration three and four, they would have included an "assessment" iteration. This would have been a time to inventory progress—after three iterations and many deliverables—and to map out next steps in the context of what had been accomplished to that point. Staff contend that they did not build ample time into the schedule for thorough assessments during implementation. The schedule was tight; however, they were able to keep to it. The tight schedule was due in part to the time it took to secure a vendor due to the state's procedural requirements for projects of this size.

Data Sources

PEAKS contains data from multiple systems. These include:

- The ACCESS system, which is the transactional system for the child support program. ACCESS contains child support case data, including collections data. In addition, ACCESS interfaces with many systems, such as the Department of Motor Vehicles.
- The state financial system, which includes program expense data.
- OCS's imaging system, which includes data about scanned images of OCS case documents and files (e.g., indicates if a case was imaged).
- Unemployment Insurance (data come from the Department of Employment and Training through ACCESS, to PEAKS).

Data are available down to the individual level.

In order to determine the elements that would be included in PEAKS, OCS conducted a requirements analysis. DSS staff met with each OCS function group to determine their data needs. OCS then assembled a list of potential data elements to be included in the data warehouse. The final elements were selected from this list.

There is capacity to add new elements to the system. OCS is planning to include additional data related to payroll and OCS personnel (the Human Resources Management System). In the future, additional data from ACCESS (Food Stamps, Medicaid, and TANF data), the Departments of Employment and Training (wage data), Corrections, Social and Rehabilitation Services, the Credit Bureau, Vermont Family Court, and the Social Security Administration will be included. The data will be used to study the child support program's influence on other programs, matches with wage records will determine sources of income, and matches with court records will indicate the progress of cases and cycle time. OCS would like to include court data (e.g., number of magistrate hearings, time cases spend in court or "cycle time") that would both help staff track cases from beginning to end and substantiate the amount of time that courts spend on IV-D business (for reimbursement purposes).

Data Uses

Federal Reporting	No, but used to check data as a means of quality control.
Performance Measurement	Yes
Self Assessments	Yes
Ad hoc Capabilities	Yes
Types of Analyses	Perspective of current status of cases and trending over time.

The primary function of the data warehouse is program management. PEAKS enables OCS program managers and caseworkers to access and analyze child support data more effectively and make informed decisions to improve program performance and enhance child support services. A secondary function is policy development.

Program management and policy tools include a "Digital Dashboard," pre-defined reports, ad hoc capabilities, and data mining. Each is described below.

*Digital Dashboard***.** The digital dashboard includes summary information on performance. One display depicts monthly collections compared with the state's collection goal and prior year collections. Another shows the support order percentage (the trend over the year and by region). The human resources section lists the number of vacant positions. Data are refreshed daily. The Dashboard is located on the user home page.

*Pre-defined reports***.** Users can select a pre-defined document on the home page. Examples of pre-defined reports include:

- Cases to review for closure.
- Potential audit error case list.
- Cases with arrears but not receiving payments.
- Children in need of paternity action.
- Number of actions taken by workers (workers wanted this to demonstrate how busy they are).
- Arrears only cases with low arrears (used to contact parties to prompt settlements of arrears).

OCS developed these reports based on previous requirements sessions. The content of the reports demonstrated that OCS was willing to provide staff with relevant information based on their needs. OCS continually gets feedback on the reports from staff and assesses whether they are intuitive to all staff.

Federal reports. Although the data warehouse is not used to create reports related to the Federal incentive payments (i.e., OCSE 157, 34a, 396), performance on four of the five incentive measures (paternity establishment, cases with orders, current collections, payments toward arrears) can be tracked. The OCSE 157 report is available to workers via the data warehouse site and is updated monthly. There are four levels of detail available: the state overview, the region, the case worker, and the case. For example, regional managers can continually check the progress of their regions in the four performance areas. A case worker can check his or her performance on a particular measure.

Self-assessment reports. The data warehouse is used to create annual self-assessment reports.

Ad hoc reports. Users can create their own ad hoc reports. The OLAP capability allows users to extract information and manipulate it through graphs, tables, and matrices. A user can drill into a report to uncover details that comprise the summary information. For example, a user could select from case data, case financial data, and person (custodial parent, non-custodial parent, children) at different levels of detail (e.g., state, region). A cross-tab table could explore case type by region and total received during the month. Other examples of possible ad hoc reports include:

- Customer comment card reports (feedback from customers).
- Cases in Locate with a collection.
- Number of cases, by type, on an individual worker's caseload.
- Collections in a specific region or county by worker.

Many users retain their own ad hoc reports that are not published to the general user community.

Policy research. OCS uses the data warehouse to analyze patterns and trends. For example, the agency used the data warehouse to explore factors associated with stoppages in payments among non-custodial parents who had previous payment histories to determine if actions could be taken to prevent stoppages. (The analysis found that "stoppers" were likely to be young—19 to 23—newly married, associated with a young custodial parent, and cases where reduced OCS contact occurred).

OCS also used the data warehouse to explore the accuracy of paternity data reported to OCSE. Staff used the data warehouse to predict, based on case information, whether a child was likely to have an error in the paternity code. Staff made 1,300 changes to paternity data based on the analysis prior to the Federal reporting deadline.

In addition, OCS explored cases that suddenly began making payments to determine what precipitated the change in status and the possible actions OCS could take to encourage the change in status. The strongest correlation was level of communications with the non-custodial parent. That is, more OCS communication was related to a higher likelihood of payment. OCS also might explore cases with frequent contact from the customer to determine what types of services could be offered that would result in a reduction of calls to OCS. Additionally, OCS would like to use the data warehouse to explore the impact of child support collections on TANF exit.

Data Warehouse Users

All OCS employees have access to the data warehouse. Other state agencies cannot access the data warehouse, nor can external entities (e.g., university researchers).

OCS anticipated three groups of users: Power users (i.e., those who would create many ad hoc queries), moderate users (i.e., those who would use the pre-defined reports and maybe add some additional fields), and casual users (i.e., those who would view the Dashboard and pre-defined reports). Currently, case workers and supervisors have access to the same level of data. When personnel information is integrated into the data warehouse, some data will be restricted. Users access the data warehouse through the network. There is a single sign-on, which needs authentication.

OCS provides training through a "train-the-trainer" approach. The power users were trained first with the expectation that they would embrace the new technology and quickly familiarize themselves with the data warehouse's capabilities. Supervisors were included in this group. OCS wanted to get the buy-in of the power users first. The power users train others. Currently, there are three types of training: power users, technical training and one-on-one training.

After training the power users, OCS rolled out the data warehouse at a supervisors meeting and walked through the newly developed reports available.

Funding

The primary funding sources for the data warehouse were the OCSE 1115 grant (29 percent), the state's share (5 percent), and the Federal Financial Participation (FFP) match (66 percent). The cost to build the data warehouse was about $2 million over 3 years.

The state has yet to fully determine the maintenance costs, in part because it is still in the initial development phase. The business case estimated that $17,000 (from state general funds) annually would be necessary for maintenance and operations (with a $33,000 FFP match). However, OCS hired an additional staff person who would be responsible for management and programming functions. In-house staff are being trained to complete future upgrades to Business Objects. OCS might also hire a contractor to help with data mining activities.

Staff anticipate few costs associated with data expansion as they hope to do much of this work in-house.

Architecture

Design Issues	
How often are data collected/corrected?	Weekly, moving to daily
How are data linked?	Data are linked in the child support source system
How many tables are in final database design?	174
How much data are retained?	Entire case history since 1981; 26 gigabytes over the life of the system
What are the security requirements?	User accounts and passwords for web-based access. Currently all users have access to same level of data; this will change when personnel information is integrated into the warehouse.
Data Source Descriptions	
Hardware platform	Dell 1650 Server
Operating system	Winders 2000
Database or file type	SQL Server
Extraction, Transformation and Loading	
Number of source systems	Four
Total data transported	A few hundred megabytes per month
End User Access	
Number of business areas in system	Eight
Tools used	Business Objects, Web Intelligence, Digital Dashboard, DT

WASHINGTON CASE STUDY

Data Warehouse Snapshot

Date created	1997 (DSS); 2001-2004 (data warehouse)
Primary uses	Program management, policy development
Primary data sources	Child support automated system (SEMS), Employment Security Division, TANF
Types of reports	Primarily ad hoc
Users	Child support staff
Primary funding source(s)	1115 grant (data warehouse); child support agency funding and 1115 grants (research arrears; orders) (DSS)

Background

The Management and Audit Program Statistics (MAPS) unit within the Division of Child Support (DCS) conducts policy development research and statistical analyses to understand the consequences of DCS actions and policies on outcomes for the agency and its customers.

MAPS designed and implemented an information system housed on its intranet called the Decision Support System (DSS), which has been active since 1997. DSS is a "point and click" system that enables IV-D staff to access data, draw random samples, and formulate ad hoc inquiries down to the individual caseworker level.

In 2001, MAPS received an OCSE 1115 grant to fund the development of a data mining system. MAPS brought historical records and multiple databases under one structure, the DCS Data warehouse.[15] Specifically, the data warehouse's data mining capabilities will enable the creation of statistical models that can be used for cause and effect analysis (e.g., how support enforcement worker actions affect outcomes) and predictions. In addition, data mining offers DCS the opportunity to find patterns of data coding errors.

The DSS (and the data warehouse) is used for policy analysis, internal reporting, performance measurement reporting, and ad hoc queries. It is not used to create Federal reports. The child support mainframe (SEMS) creates the Federal reports per Federal auditors' instructions.[16]

Implementation

Work began on the data warehouse in 2001. The original completion date was September 2004, although MAPS is requesting a no-cost extension from OCSE through September 2005.

Much of the first year was spent acquiring the technical resources necessary to build a data warehouse with the desired mining capabilities. An internal technical network expert assisted MAPS staff in the design. Initial activities were slowed during the first three months due to budgetary issues. MAPS had to wait until the state authorized use of the 1115 funds. It took three months to resolve the issue, causing delays in purchasing, scheduling training, and

[15] To this end, MAPS created a Windows domain for file serving, file sharing, and database applications. Within the network of servers is a Network Attached Storage (NAS) that allows large files that are continually expanding to be stored in one place.

[16] Federal audit regulations require that if errors are found in the data, the changes need to be made in the mainframe system, not the data warehouse.

subsequent work. In addition, MAPS staff had to develop an Information Technology Portfolio for the Department of Information Services (DIS) regarding security requirements. This process was time and resource intensive, but MAPS completed the portfolio by the end of the first year.

MAPS also had to overcome some initial programming inefficiencies. The first contractor built static coded business rules that would impede the addition of new elements or changes to the definitions of existing data elements. As a result, MAPS did not renew the contract. The second contractor was able to work with project staff to convert the static coded business rules into a reusable, table-driven business rule approach.

The implementation process also involved modifying and establishing data sharing agreements. In addition to child support data, the data warehouse contains data for TANF clients from the Employment Security Division (ESD) and the Department of Social and Health Services (DSHS). Data sharing agreements with each agency were in place for data use related to child support enforcement functions, but the agreements did not allow any other use. MAPS had to amend the data sharing agreements to include non-enforcement activities. Other data sharing agreements had to be developed, the most complex of which is a two-way agreement that will allow matches to non-custodial and custodial parents who are recipients of social services in addition to TANF.

Data Sources

The primary data sources are:

- SEMS—the child support mainframe—is the main source of data. It includes order, child, case, and payment files.
- ESD—wage information.
- The IV-A data warehouse. It includes TANF case characteristics (e.g., date of the welfare grant, grant amount, number of children on the grant).

MAPS staff are currently working with the DSHS research and data analysis division to gain access to the data repository that contains client service data from the nine divisions of DSHS (including Medicaid, TANF, and Food Stamps). The Client Services Data Base (CSDB) is a large relational database system that reports on services in common among the different agencies and includes a customer satisfaction survey. It can be used for geographic mapping based on addresses. The data from the CSDB will allow MAPS to quantify the costs of services provided to non-custodial parents, making it possible to better understand the types of barriers and the costs of addressing them. MAPS has completed the first phase of data exchange by sending child support data to the CSDB. The second phase will be for DSHS to send client service data to MAPS to match client records.

During implementation MAPS staff gathered as much data from the legacy system as was possible, focusing on key data elements that were critical to both internal and external reporting. The ability to add new elements to the data warehouse was an important design consideration. MAPS recently added child relationship to the non-custodial parent and the medical insurance carrier codes. In the future, MAPS may add information about arrests.

Data Uses

Federal Reporting	No
Performance Measurement	Yes
Self Assessments	Yes
Ad hoc Capabilities	Yes
Types of Analyses	Point in time, trends

Currently, the DSS functions are program management and policy development. This includes the incentive performance measures down to the individual worker and all types of data inquiries for internal and external audiences. MAPS uses the system to produce monthly and quarterly reports. There are also ad hoc capabilities. Each is described below. In addition, the DSS is used to pull a sample of cases for the annual Federal self-assessment and for the internal data reliability audit, which mirrors the Federal audit.

Reports. The MAPS unit provides monthly "progress reports" on Federal and state performance measures, to the individual worker level. A number of reports are on the web (static and historical). Most reports are updated monthly, including: Governor's reports, Federal incentive reports, collections/establishment reports, retained support report, performance measurement matrix, retained support sorted by field office/team/state, overview of performance measures, and tribal Federal performance measures. Roll-up reports are less frequent (quarterly).

Examples of reports include:

- Retained support. A user can look at monthly goal versus actual performance, as well as IRS money versus non-IRS money by team and support enforcement officer within the team.

- Overview of performance measures. Users can select performance by support enforcement officer number and date (e.g., March 2002 through March 2004) and can table and graph the results. Also, users can compare workers.

- Performance report. This reports on the Federal incentive measures and its components and has drill-down capability from the state to the field office, to a team, and ultimately to an individual case worker. In addition, graphing capabilities with some choices for appearance can be created with point-and-click ease. Findings indicate areas where training would be appropriate.

Ad hoc capabilities. Staff can use the DSS to create their own reports. This is a point-and-click system that allows staff to access data and formulate ad hoc inquires at the state, field office, team, or case worker level. Staff can generate random samples, download data into spreadsheets (e.g., Excel) or databases (e.g., ACCESS) and get counts of cases meeting their criteria. Users can select variables (e.g., enforcement, order data, case type, individual information) by field office and support enforcement worker. For example, a user might want to know how many of a particular team's cases had no order. Or, a worker can see his or her performance on four of the five Federal incentive measures (excluding cost-effectiveness). Data can be displayed in tables and/or in graphs for a set period in time or over time.

Policy research. The 1115 grant is being used to create data mining capabilities that can be used to create statistical models that allow predictability, cause and effect, and the development of an outcome-based system. Along these lines, MAPS has conducted a number of studies, including:

- Cost avoidance. Staff found that regular child support payments produce savings to the TANF, Medicaid, and Food Stamps programs. For example, previous research showed that custodial parents who received regular child support payments subsequently used less welfare and had more employment. These studies were based on custodial parents who were on welfare during a specified time. With the data mining capability, MAPS expanded on this previous work by considering all identifiable custodial parents in DCS records from January 1998 forward (417,870 parents). The analysis found that at each wage level across 59 months of data, custodial parents who received regular child support used 50 percent to 75 percent less welfare than custodial parents who did not receive regular payments.
- Arrearages. Staff found that if the monthly order is greater than 20 percent of earnings, the non-custodial parent is more likely to build debt.

MAPS staff will use the data warehouse to learn more about causal relationships between individual characteristics, timing of life events, and the payment of child support. For example, through the analysis of orders, the state's support schedule will provide the context for understanding the relevance of order amounts, especially the consequences of default orders and imputed income. An in-depth analysis on hard-to-collect cases will document why parents do not support their children and what collection techniques work.

MAPS audit staff also use the data warehouse to conduct internal data reliability reviews. The data mining function is used to find patterns of coding errors in SEMS. This information is vital to the Audit Team who will be working closely with the Training Team to develop specific case-coding training in a variety of ways, including on-line training, to eliminate errors at the source.

Data Warehouse Users

Only MAPS staff have direct access to the data warehouse and individual-level data. DCS staff (including line workers) have indirect access through the DSS application (i.e., they can view and manipulate data, but not alter it). The DSS was built for field office staff to help analyze and work cases. In the DSS, performance data can be drilled down to field offices, teams and individual workers. Workers within DCS' firewall can ask for data down to the IV-D number, but not names or Social Security Numbers.

DCS staff access data through a web-enabled interface. The system largely is self-explanatory. There is an online tutorial for the DSS. MAPS tries to educate staff so they can run their own reports. Information can be downloaded into Excel easily (click a button).

At this time, few non-IV-D staff have access to the data. Some prosecutors can view the data (if they are authorized to use SEMS), as well as some tribal offices. IV-A does not have access, but MAPS is considering creating an application for them such as a "digital dashboard" that has summary information and no individual-level data.

Funding

The cost of building the data warehouse was $2.1 million, funded through the 1115 grant and the state match. In addition, the state pays for the staff's time (the data warehouse was developed internally).

Maintenance costs are unknown at this time. However, the primary costs will be for software licensing and upgrades. The only cost for data expansion would be staff time.

Funding for the development of the DSS came from the child support agency and 1115 grants.

Architecture

Design Issues	
How often are data collected/corrected?	Monthly, except for employment data, which is quarterly. Data are not corrected; it is taken as is.
How are data linked?	Data are linked by the case number assigned.
How many tables are in final database design?	11
How much data are retained?	From 1997 forward, everything is retained.
What are the security requirements?	IIS (Internet Information Services, web server) usage tracking. Only MAPS staff have access to individual-level data; others have view-only access.
Data Source Descriptions	
Hardware platform	Intel x86
Operating system	Windows 2000 Advanced Server
Database or file type	MS SQL Server
Extraction, Transformation and Loading	
Number of source systems	20
Total data transported	Not yet tracked
End User Access	
Number of business areas in system	Two: MAPS staff has direct access and the Division of Child Support has indirect access
Tools used	Six

WISCONSIN CASE STUDY

Data Warehouse Snapshot

Date created	2002
Primary uses	Policy development
Primary data sources	Child support automated, wage data, TANF data
Types of reports	Ad hoc
Users	Child support staff; Institute for Research on Poverty (University of Wisconsin)
Primary funding source(s)	1115 grant, FFP

Background

Wisconsin has multiple data warehouses that are program-focused (TANF, Food Stamps, Medicaid, child care, and earnings). In 2002, the Department of Workforce Development (DWD) Bureau of Child Support (BCS) received an 1115 grant from OCSE to incorporate child support payment and case data of work program individuals into the existing earnings data warehouse, which since 1997 has captured wages of TANF recipients one year prior to entering the program and three years following program exit.

The new earnings/child support data warehouse ("the data warehouse") was created so that BCS staff and researchers could efficiently access and analyze aggregate data. The data warehouse is also used for program management.

The data warehouse can be used to track outcomes and make business cases for operating programs. For example, two state employment programs work with unemployed or under-employed non-custodial parents (Children First and the Workforce Attachment and Advancement Program).[17] In response to a request from the state legislature, DWD conducted studies that found an increase in earnings and child support payments following participation in work programs. The state's new administration is interested in building on these work programs.

In addition to BCS, the DWD Bureau of Workforce Information and the University of Wisconsin, Madison, Institute for Research on Poverty (IRP), were involved in the decision to create the data warehouse. The ability to study interrelationships among the welfare and child support programs was a powerful motivator for all entities.

Implementation

The project began in April 2002 with planning sessions, which included staff from BCS, IRP, the Bureau of Workforce Information, and the Bureau of Information Technology Services. These sessions focused on the business questions the system would be able to answer. Then technical staff researched feasibility of all data needs. Development of the system began in August 2002. The system was developed and tested by December 2003.

[17] The Children First program, a court-ordered program for under- or unemployed non-custodial parents, is funded by state general-purpose revenue and used as maintenance-of-effort funds for the TANF program. WAA provided training, job retention, basic skills development and support services to TANF-eligible custodial and non-custodial parents. It was funded for two years through the TANF block grant.

The timetable was extended due to problems encountered with loading historical KIDS (child support) case information. At the time the grant was applied for it was unknown that KIDS did not keep historical case information but overwrites that information as new entries are made. This setback was unexpected and took many hours of research to determine if there were solutions. It was decided to continue loading data from the most current data available rather than trying to get incomplete historical KIDS case data.

In hindsight, staff would have made a few changes to the current design of the data warehouse in order to make it more useful. For example, staff continue to discuss adding more data elements and/or adding all child support cases into the data warehouse to increase policy development and program management capabilities. Staff believe that the earnings data warehouse was the right one to attach the child support data to.

Data Sources

The data contained in the data warehouse include:

- Child support case and individual information from KIDS (Wisconsin's automated child support system). The child support data in the data warehouse is a subset of the KIDS data (child support cases that are currently in or have been in a work program). Child support data elements include custody status, child support paid, and child support received.
- Quarterly wage and employer data from the Unemployment Insurance database.
- Data on TANF work programs from the CARES system (the TANF, Food Stamp, and Medicaid mainframe).

The data warehouse does not interface directly with CARES and KIDS. The interface is not direct but through the use of extracts, which are point in time snapshots of the data as it existed the day the extract was run.

KIDS data is available at the individual, case, local office, county and state levels. Earnings data are available at the individual, employer, county and state level.

Discussions with technical, research, and child support program experts determined which elements would be in the system. New elements can be added. While no additions are being made at this time, data warehouse staff would like to include Department of Revenue data (e.g., the EITC, household income, and wages from out of state) in the future. Currently the data warehouse has data on earnings but not other income sources.

Data Uses

Federal Reporting	No
Performance Measurement	No
Self Assessments	No
Ad hoc Capabilities	Yes
Types of Analyses	Point in time, trends, projections

The data warehouse is used for policy and program development; thus, the intent of reports is different than those developed for other systems (e.g., KIDS). The primary function of the data warehouse is to explore interrelationships and impacts of Wisconsin's child support and work

programs, for both custodial and non-custodial parents. Types of analyses include point in time, trends, and projections.

Reports. Reports are ad hoc in nature but users have the capability to create summary level data at the individual, case, office, county, or state level. Monthly child support data analyses are available for the individual, payment, and case. Quarterly data are available for work programs.

Data warehouse staff are currently developing prototype reports for counties with the Children First program (39 of 72 counties) to answer program management questions. These include the extent to which participants in work programs are employed before and after the program; average earnings before and after the program; child support orders before and after the program; and child support payments before and after the program.

Local officials and case workers cannot use the data warehouse to track performance on incentive measures because not all child support cases are in the data warehouse. However, the data warehouse can be used to track county performance on the Children First program. Findings can be used to clarify policies, if necessary.

Examples of analytic and policy questions that are answered by the data warehouse include:

- Does a non-custodial parents' participation in work programs increase or improve employment? Does increased/improved employment result in increased child support payments?
- Do non-custodial parent participants in work programs exhibit earnings increases over time? If so, at what rate? Do the earnings increases result in higher child support payments? If so, is there a lag between increased earnings and increased child support payments?
- Does the receipt of increased child support by the custodial parent change the custodial parent's participation in the labor force? Does the non-custodial parent increase child support payments over time, when working and paying a "fair share" of support for the child(ren)?
- What share of child support orders might be considered for modification as a result of increased custodial or non-custodial parents' earnings due to work program participation?
- Are there geographic variations in child support payments for non-custodial parents who participate in work programs? Are there geographic variations in custodial parents' earnings due to work program participation?
- Does it appear that the availability of work programs encourage non-custodial parents to establish paternity? Does this vary by geography?
- Do orders to participate in Children First result in immediate payments of arrears? At what rate? How many non-custodial parents ordered to participate in Children First actually participate in job search or other aspects of the Children First program?
- What types of employment are non-custodial parent participants in work programs obtaining? Custodial parent participants?

- Are there any differences in terms of changes in earnings or child support paid among demographically similar cases in which both parents participate in work programs and cases in which only one parent participates in work programs?

Data Warehouse Users

Access to the data warehouse is limited to staff from the BCS, the Bureau of Workforce Information, and three IRP staff.

There are different levels of access. A "viewer" can see data but cannot create separate reports; an "editor" can edit previously created reports but cannot create a new report; and an "expert" can create a report using any of the measures and dimensions. Experts can create their own queries. IRP staff are the expert users. Their access is not restricted because they have a data sharing agreement in place. If no data sharing agreement exists, aggregate data can be viewed but not individual or case specifics.

Data are accessed through Web Intelligence, a Web-enabled interface, or ad hoc report requests. Anyone can request a report but the level of detail available depends on the type of end user. Reports are requested by calling DWD and requesting data. There are not many requests for the data from outside of DWD and IRP, in part because the data warehouse is not advertised (there are sensitivities around wage data). Web Intelligence training is provided by DWD technical staff.

Funding

BCS has spent $446,674 on development of the data warehouse. The primary funding sources are child support FFP, Federal Indirect Cost Reimbursement (FICR)[18] and the 1115 grant. The cost of maintenance is expected to be similar to those of the state's other data warehouses ($30,000 per year). This includes 120 hours of batch support, Oracle storage, and license for Web Intelligence. It does not include the cost of adding additional data elements. This cost is minimal unless an entire universe needs to be created, which costs approximately $65,000.

[18] State services that they cannot charge the federal government for (e.g., employee relations)—the rate is applied to federal grants. The money becomes state funds that can be used as a match.

Architecture

Design Issues	
How often are data collected/corrected?	Monthly from the child support source system and quarterly from the eligibility system mainframe and unemployment system. Unemployment earnings reports are corrected every quarter for the previous four quarters based on revised unemployment earnings information.
How are data linked?	Data are linked in the child support source system.
How many tables are in final database design?	60
How much data re retained?	All extracts are maintained on the mainframe; databases are retained indefinitely staging tables are truncated monthly
What are the security requirements?	Security handled through Oracle; front end OLAP security is set up in Business Objects and relies on individual access rights for their particular profile. IDs granted only to those who complete the proper paperwork to indicate their level of clearance. There are also different levels of access. Experts have access to individual-level data because they sign data sharing agreements. Others have access to aggregate-level data.
Data Source Descriptions	
Hardware platform	
Operating system	
Database or file type	IBM DB/2 (mainframe)
Extraction, Transformation and Loading	
Number of source systems	Three
Total data transported	0.6 gigabytes per month
End User Access	
Number of business areas in system	Three: Bureau of Child Support, Bureau of Workforce, Institute for Research on Poverty (University of Wisconsin)
Tools used	Web Intelligence, Business Objects

WYOMING CASE STUDY

Data Warehouse Snapshot

Date created	In process
Primary uses	Medical support enforcement, performance management
Primary data sources	Child support automated system (POSSE)
Types of reports	Pre-defined
Users	State IV-D staff, district managers, limited court staff
Primary funding source(s)	1115 grant

Background

The Wyoming Child Support Enforcement Program received a three-year OCSE 1115 grant in 2002 to build a data warehouse. The data warehouse is currently under development.

Child support staff was integral to the creation of the data warehouse. However, there was a lengthy approval process that involved state committees that oversee agencies, the state Chief Information Officer, the state purchasing office (which must approve the RFP for a vendor and the resulting contract) and the Attorney General's office. The project needed to be "blessed" at every level. The project did not encounter many difficulties during these various reviews, but the process was time consuming.

Implementation

In addition to the approval process, the child support agency experienced an issue with IRS. The IRS stated that contractors cannot have access to IRS data. Wyoming, however, needed contractors to build the data warehouse, so the child support agencies granted them access. Wyoming's IV-D program, along with some other states, are challenging the IRS' policy. According to child support staff, the IRS and OCSE are discussing the issue.

Data Sources

The primary data warehouse data source is the automated statewide child support system, Parental Obligations System for Support Enforcement (POSSE). POSSE currently interfaces with the following systems:

- TANF
- State and Federal new hire directories
- Federal Case Registry
- State Department of Labor
- Financial Institution Data Match
- Bureau of Game and Fish
- Department of Motor Vehicles
- Department of Workers Compensation
- Social Security Administration

There was limited medical information in POSSE. With the 1115, the child support agency is creating interfaces with the Medical Management Information System (Medicaid) and Blue Cross (operator of the SCHIP program).

The data warehouse data elements were guided by the 1115 grant purpose (medical support enforcement). However, child support staff also needed specifics about cases (e.g., case information, person information, locate status). There were numerous meetings between child support staff (specifically the end users) and the vendor. The users prioritized a list of 20 pre-defined reports; the data support these reports. Data are available at the district office, individual worker, and case level.

Not all desired elements were included. According to child support staff, POSSE is not user friendly. Staff wanted to crease a history of each case that could easily be loaded into the data warehouse. This proved to be very difficult. Thus, there will not be histories of each case since inception; rather, information is limited to the date the data warehouse went live (July 31, 2004) and forward.

There is capacity to add new elements to the system. However, new elements are not being considered at this time. The staff will assess how the data warehouse performs when it goes live in August. Eventually, the state would like to include an interface with WYCAPS (child care and child protective services) and IRIS (the new TANF information system that will replace the current system, EPICS).

Data Uses

Federal Reporting	Yes (data accuracy testing)
Performance Measurement	Yes
Self Assessments	No
Ad hoc Capabilities	In future
Types of Analyses	Point in time, point in time comparison (going forward), trends (going forward), projections

The primary function of the data warehouse will be tracking and enforcing medical support orders. The data warehouse will provide a clearer picture of medical support in Wyoming, including the number of children with insurance, the number of children receiving public health benefits (including SCHIP), and the number of employers offering medical insurance benefits. Child support staff will also study how these statistics compare with the actual court orders for specific cases (e.g., is the child receiving coverage as ordered?).

The data warehouse will also include a multitude of child support data, so it will be used to track performance of local (district) offices. District offices are operated by vendors under contract to the state IV-D program. These are performance-based contracts (i.e., performance is based on the five Federal incentive measures). The data warehouse will be used as a management tool to help district offices identify areas where improvement is needed.

Finally, the data warehouse will be used to examine the data that are used to calculate the 157 report line-items for accuracy and integrity.

At this time, the data warehouse will produce pre-defined reports. Ad hoc capabilities and data mining functions will be added at a later date.

Pre-defined Reports. The data warehouse will produce 20 pre-defined reports.[19] The reports will be updated monthly. The types of analyses available include point in time, point in time comparison (going forward, but currently there is no historical data in the data warehouse), and trends (going forward). Forecasting reports (for collections and income withholding) are used for strategic planning. Reports that track enforcement activities (e.g., if more money is spent on revoking fishing licenses, will collections improve?) also contributes to strategic planning.

Ad hoc reports. Ad hoc reports are not available at this time. The state will begin taking requests in early 2005. The data warehouse ad hoc report requests will be handled in the same way as requests for POSSE reports: district managers or state-level supervisors can go on-line and register a request. Technical staff will prioritize the requests.

Data Warehouse Users

There were 24 users when the data warehouse went live. This includes nine district managers, one court official (for a total of 10 users in the field), "super users" in the state office, programmers, and limited state-level staff.

The super users have access to all 20 pre-defined reports, in addition to ad hoc capabilities. The district managers in the field have access to nine of the pre-defined reports that can be used to manage performance. The court official has access to one report. Some reports are limited to specific state-level specialists (e.g., medical reports, forecasting reports, incentive reports). In all of the reports, individual-level data will be displayed but cannot be changed. All staff sign confidentiality agreements. District managers can only see data for their own district. The pre-defined reports will be available on the users' desk tops.

Users are trained through a vendor-produced manual and hands-on-training (i.e., learning by doing). In addition, data warehouse developers met with district managers and state-level staff and showed them how to access and use the pre-defined reports.

Staff from other state agencies and other entities (e.g., universities) will not have access to the data warehouse.

Funding

The primary funding source for the data warehouse was the OCSE 1115 grant and the 5 percent state match. The initial start-up cost was $1 million over a three-year period.

The maintenance costs have not yet been determined, nor have the costs associated with data expansions.

[19] Arrears Cases With No Payments; Asset Report; Case Closure Cleanup Report; Cases Paying Towards Current Not Arrears; Collection Forecast; Collections By County By Enforcement Method; Collections By District By Enforcement Method; Collections By State By Enforcement Method; Demographics Reports (for both custodial parents and non-custodial parents, broken out by disability, employment, incarceration and marital status); Federal Judgment Collection Report; Initiating Interstate Case Report; Intake Time Frame Report; Interstate FIPS Code Report; Interstate Report; Medical Information Detail Report; Medical Summary; Non-custodial Parent Located No Payment Previous Calendar Month; Non-custodial Parent Not Located Report; Non-IVD Case Report; Withholding Forecast.

Architecture

Design Issues	
How often are data collected/corrected?	Daily
How are data linked?	Data are linked in the child support source system (POSSE)
How many tables are in final database design?	34
How much data are retained?	At this time, there are no historical data; data will be retained going forward
What are the security requirements?	Only "super users" have access to individual-level data. Others access predefined reports.
Data Source Descriptions	
Hardware platform	Dell 4 Server
Operating system	Windows 2000
Database or file type	SQL Server
Extraction, Transformation and Loading	
Number of source systems	42
Total data transported	0.42 gigabytes daily; 13 gigabytes initially loaded
End User Access	
Number of business areas in system	Child support
Tools used	Esperon