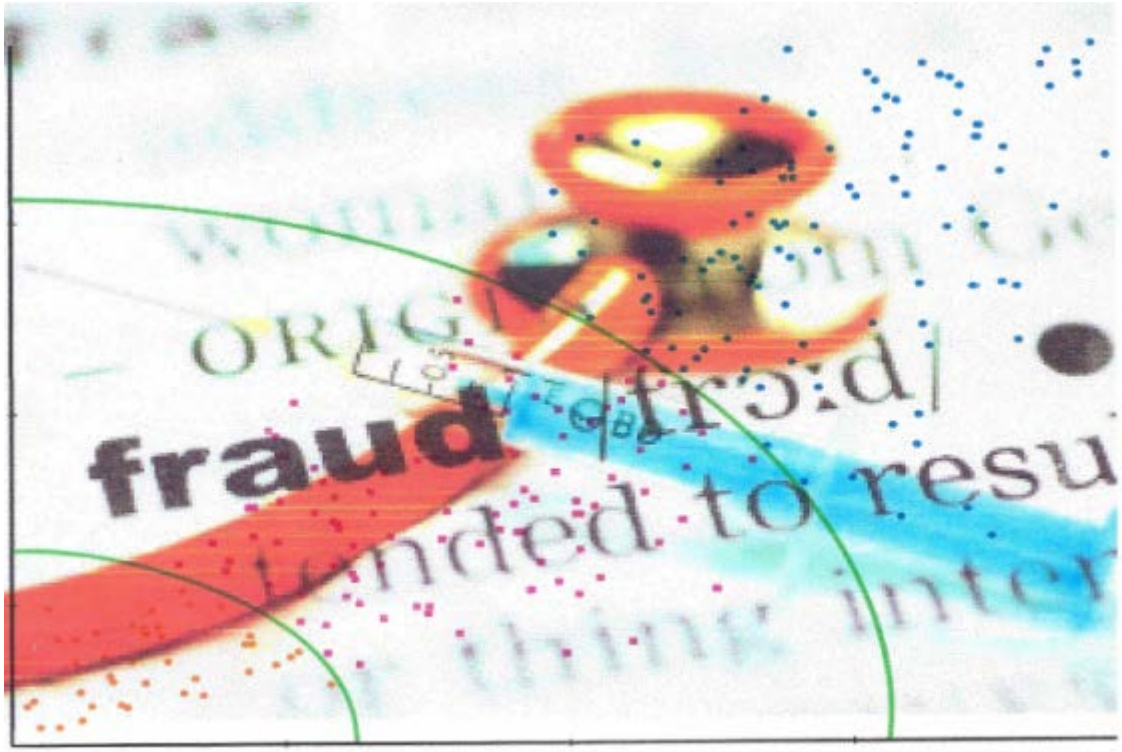
FINAL MEDS AD Waiver Evaluation: Data Mining Activities Interim Report Preliminary Findings

Prepared for
Florida Medicaid MED 143

Project 2, Deliverable 9



College of Medicine
Florida State University
June 27, 2013

Table of Contents:

1. Background and Perspective 7

2. Data Mining Activities Statistics 14

   2.1. Input: Budget, FTEs, and Training 15

   2.2. Output: Complaints, Opened New Cases, Cases Investigated, Disposition of Cases 22

   2.3. Outcomes: Monies Recovered 32

3. Data Mining Activities Key Informant Experiences 36

4. Data Mining Activities Preliminary Evaluation 45

5. Preliminary Conclusion 52

   Appendix: Production Function Used 54

List of Figures:

Figure 1: Structure-Conduct-Performance-Paradigm (SCPP) transposed on MFCU/DMI, AHCA and Other State and Federal Agencies.

Figure 2: Input – Throughput – Output – Outcome Model.

Figure 3: MFCU Budget, MFCU Grant and Data Mining Grant (Federal and State Matching Funds), FFY 2006-07 through FFY 2012-13.

Figure 3a: MFCU Data Mining Initiative (DMI) Budget, Federal Data Mining Grant and Florida State Matching Funds, FFY 2010-11 through FFY 2012-13.

Figure 4: MFCU Budget and Expenditures, MFCU Grant and Data Mining Grant, FFY 2006-07 through FFY 2012-13 (YTD).

Figure 4a: MFCU Data Mining Initiative Budget and Expenditures, Federal Grant and Florida State Matching Funds, FFY 2010-11 through FFY 2012-13 (YTD).

Figure 5: "MFCU-Opened" New Cases out of all Complaints, FFY 2010-11 through FFY 2011-12.

Figure 6: Relative Shares of Opened New Cases by Source, FFY 2006-2007 through FFY 2012-13 (YTD).

Figure 7: Total Amount of Monies Recovered by MFCU, FFY 2001-02 through FFY 2011-12.

Figure 8: Number of Cases Investigated Relative to the Total Amount of Monies Recovered in Millions, Average SFY 2006-10, SFY 2010-11 and SFY 2011-12.

Figure 9: Total Amounts of Monies Recovered and Total FFP + Florida, SFY 2007-08 through SFY 2011-12.

Figure 10: Various Tabs of an Investigative Data Mining Activities Report.

Figure 11: Number of Complaints, Opened New Cases, Disposition of Cases, and Cases Ending in Settlement, Conviction, or Plea Agreement, MFCU, FFY 2010-11 and FFY 2011-12.

Figure 12: Number of Complaints, Opened New Cases, Disposition of Cases, and Cases Ending in Settlement, Conviction, or Plea Agreement, Attributed to DMI, FFY 2010-11 and FFY 2011-12.

Figure 13: Actual versus Expected Number of Opened New Cases MFCU, FFY 2006-07 through FFY 2012-13 (YTD).

Figure 14: Sensitivity Analyses of Average Budget and Full Time Equivalent Employment on Expected Number of Cases.

List of Tables:

Table 1: MFCU Full Time Equivalent (FTE) Employment incl. Data Mining Analysts, Budgeted versus Applied, FFY 2006-07 through FFY 2012-13.

Table 1a: MFCU Full Time Equivalent (FTE) Data Mining Analysts and Approximate Hours Devoted to Data Mining, per MFCU Region, FFY 2006-07 through FFY 2012-13.

Table 2: Top Six Course Titles in Time Allocation for Training of MFCU Data Mining Analysts, FFY 2011-12 and FFY 2012-13 (YTD).

Table 3: The Number of all Fraud Complaints Received by the MFCU, FFY 2006-07 through FFY 20120-13 (YTD).

Table 4a: The Top Eight Sources by Number of all Fraud Complaints Received by the MFCU, Broken Down by Source, FFY 2010-11 through FFY 2012-13 (YTD).

Table 4: The Number of all Fraud Complaints Received by the MFCU, Broken Down by Source, FFY 2010-11 through FFY 2012-13 (YTD).

Table 5: Top Five Number of all Fraud Complaints Received by the MFCU, Ranked by Provider, FFY 2010-11 through FFY 2012-13 (YTD).

Table 6: Number of Fraud Complaints Received by MFCU, by Provider Type, where the Source was Data Mining Initiative, FFY 2010-11 through FFY 2012-13 (YTD).

Table 7: MFCU Cases Investigated, Cases Opened, and the Source of the Cases, FFY 2006-07 through FFY 2012-13 (YTD).

Table 8: Opened New Cases by Region; DMI and Other Sources, FFY 2010-11 through FFY 2012-13 (YTD).

Table 9: Top Five of Medicaid Fraud Cases by Provider Type, FFY 2010-11 through FFY 2012-13.

Table 10: Disposition of MFCU Cases Closed and Subset of Cases Closed Attributed to the Data Mining Initiative, FFY 2010-11 through FFY 2012-13 (YTD).

List of acronyms

AHCA = Agency for Health Care Administration

CCEB = Complex Civil Enforcement Bureau

CFR = Code of Federal Regulations

DMAR = Data Mining Analyst Report

DMG = Data Mining Grant

DMI = Data Mining initiative

DOH = Department of Health

DSS = Decision Support System

FDLE = Florida Department of Law Enforcement

FFP = Federal Financial Participation

FFY = Federal Fiscal year

FL.AG = Florida Attorney General

FL.GR = Florida General Revenue/Program Income

FLEAT = Florida Law Enforcement Analyst Training

FTE = Full Time Equivalent

MEDS-AD = Medicaid Medications for Aged and Disabled

MFCU = Medicaid Fraud Control Unit

MOU = Memorandum of Understanding

MPI = Medicaid Program Integrity

SCPP = Structure-Conduct-Performance-Paradigm

SFY = State Fiscal year

YTD = year to date

1. Background and Perspective

Expenditures for the Florida Medicaid Program exceeded $18 billion for services rendered between July 1, 2011 and June 30, 2012. While the vast majority of those expenditures were for needed services, some of the expenditures were the result of fraudulent or abusive billing.

Fraud can be defined as: A knowing or intentional deception or misrepresentation made by a person with the knowledge that the deception could result in some unauthorized benefit to oneself or some other person.

Abuse can be defined as: Provider practices that are inconsistent with generally accepted business or medical practices and that result in an unnecessary cost to the Medicaid program or in reimbursement for goods or services that are not medically necessary or that fail to meet professionally recognized standards for health care.

In Florida, the investigation of suspected Medicaid fraud is under the auspices of the Florida Attorney General (FL.AG) at its Medicaid Fraud Control Unit (MFCU), while cases of suspected abuse are handled by the Bureau of Medicaid Program Integrity (MPI),[1] located in the Office of the Inspector General of the Florida Agency for Health Care Administration (AHCA). Staffers from AHCA, MFCU, and the Department of Health (DOH) meet regularly to discuss major issues, strategies, joint projects and other matters concerning health care.

Suspected fraudulent billing practices can be discovered in many ways, one of which is analysis of claims Medicaid has paid using AHCA's Decision Support System (DSS), which is a subset of

[1] Authorized by Section 409.913, Florida Statutes, MPI audits and investigates providers suspected of overbilling or defrauding Florida's Medicaid program, recovers overpayments, issues administrative sanctions and refers cases of suspected fraud for criminal investigation.

the Medicaid Management Information System claims database. Data mining is usually perceived as an extension of traditional data analyses and statistical approaches, incorporating analytical techniques drawn from a range of disciplines. It is important to note that data mining in itself is only a tool, since it does not eliminate the need to know the business, to understand the data, and the analytical methods involved, nor does it indicate a value to the results of the analyses. Therefore, data mining results always need translation into meaningful information. In essence there are two types or approaches in data mining; namely, approaches in which data is analyzed based on overall patterns or structure, and approaches seeking to identify departures from a norm or detect unusual data patterns. To locate these overall or specific patterns, often instructions (decision rules) or algorithms are used. There are many data-mining methodologies,[2] and all involve an assessment or evaluation of the specific approach used.[3]

As the designated "single-state-agency," AHCA's data mining activities are supported by federal funding through the Federal Financial Participation (FFP) program. Federal Financial Participation, however, has not been available to support data mining activities of staff at the Florida Attorney General's Office. The Attorney General's Office and AHCA jointly requested that this prohibition[4] be waived. On July 15, 2010, the Centers for Medicare and Medicaid Services granted a waiver of CFR 1007.19.

The Florida Medicaid Medications for Aged and Disabled (MEDS-AD) demonstration waiver provides Medicaid coverage for aged or disabled residents of the State of Florida with incomes at or below 88 percent of the federal poverty level and assets at or below $5,000 for an individual (or $6,000 for a couple). As a result of the waiver of CFR 1007.19, the MEDS-AD waiver was amended to include activities related to data mining. In particular, the amendment states:

[2] Such as SEMMA for SAS and CRISP-DM for SPSS.
[3] For further reading reference is made to J. Jackson: Data Mining: A Conceptual Overview, Communications of the Association for Information Systems (Volume 8, 2002) 267-296, and
Chung H.M.l. and P. Gray, "Current Issues in Data Mining," *Journal of Management Information Systems*, forthcoming. http://www.csulb.edu/~imats/hmchung/rp1.htm
[4] Found in Code of Federal Regulations (CFR) 1007.19

Florida Statutes § 409.913(1)

The evaluation of the MEDS-AD will be revised to include tracking of costs of data mining activities and the related recoveries or measurable cost avoidance directly attributable to analysis performed by MFCU analysts in this demonstration.

The state's quarterly reporting schedule will continue, and will include the status and progress of data mining activities related to this amendment. Tracking of costs and recoveries will be submitted by the state annually within 60 days of the end of each waiver year.

On September 13, 2010, AHCA (the "Agency") and the Florida Attorney General entered into a Memorandum of Understanding (MOU) that specifies the roles and responsibilities of the two organizations relative to data mining activities. Included in the MOU are the following provisions:[5]

Coordinate all data mining activities with the Agency, prior to commencement, to ensure actions are not duplicated.

Approximately biweekly, but in no case less than monthly, designated personnel with the parties will meet in-person to discuss data mining projects.

At or before such meeting, MFCU personnel will present Agency personnel with written proposals for data mining projects by the MFCU, if any, to review whether the proposed data mining objectives duplicate existing, or recently completed, Agency data mining projects. Meetings will also provide an opportunity to interpret the outcome of data output generated by mining projects and to exchange information regarding potential projects that will enhance the productivity and efficiency of MFCU and Agency resources.

[5] MOU Section IV.A.11 and Section VI A.2 and A.3 in particular.

By approximately the next biweekly meeting, but in any case, within one month, the Agency will provide the MFCU with written verification whether the MFCU's data mining objectives are duplicative of an existing, or recently completed, Agency data mining project. The Agency may also suggest a coordinated effort between the parties with respect to proposed data mining objectives.

In October 2010, the MFCU at the Florida Attorney General's Office commenced data mining activity.

This report presents an evaluation of the MEDS-AD waiver: Data Mining Activities, contingent on the waiver of CFR 1007.19. The purpose of the evaluation is to determine if data mining activities by the Attorney General's MFCU through the MEDS-AD 1115 (a) Demonstration Waiver have resulted in the recovery of Medicaid funds that were paid as a result of fraudulent activity on the part of Medicaid providers.

A couple of considerations must be noted. First, the Data Mining Initiative (DMI) cannot be seen apart or isolated from the activities conducted within the MFCU of the Attorney General's Office, i.e. data mining reflects on the office's overall performance. In addition, given the MOU, this performance mutually reflects on both the Florida Attorney General's Office and AHCA. Although other state and federal agencies and offices may be added, the focus of this evaluation will be at the level of MFCU and the areas of understanding between the two MOU parties, this especially with respect to the waiver provision on duplication, and the opportunity to discuss, interpret and exchange information regarding potential projects that will enhance the productivity and efficiency of MFCU and AHCA's resources. Second, the timeframe for the evaluation is rather short and only covers the timeframe of October 2010 through September 2013 (YTD)[6] (i.e., FFY 2010-11 through FFY 2012-13). Given that it takes time to build legal cases, sometimes long after data mining is done, results that can be traced to MFCU data

[6] All analyses done in this deliverable are based on year-to-date data for FFY 2012-13, unless otherwise specified.

mining activities under the waiver may not be readily available as per the timeframe of evaluation. Third, MFCU activities related to physical abuse, neglect and financial exploitation (PANE) of patients residing in long-term care facilities are not considered in this evaluation, since they don't pertain to the data mining activities

Concerning the evaluation, data mining is perceived as a tool adding a dimension to the work structure within the Florida Attorney General's Office MFCU, and likewise an opportunity to add to the inter-agency activities between the Attorney General's Office, AHCA, and possibly other state and federal agencies as well. This added dimension is highly qualitative in nature, and is only measurable by derived input variables and as far as it impacts performance. Performance will be measured in terms of output (e.g., cases) and outcomes (e.g., monies recovered); especially once it can be related to the data mining activities of the MFCU, the target activities/agency of the waiver. In addition, it is incumbent on the researchers to provide recommendations on the process of data mining and possibly on the inter-agency cooperation as mentioned.

In order to provide different perspectives, various methodologies will be used for different aspects of the evaluation; ranging from comparative analyses, attendance of meetings, interviews, literature review, questionnaires, as well as a case file review to gather information and develop insights for evaluation purposes. With respect to the evaluation, the question is:

> ***Did the Data Mining Initiative (DMI) at the Medicaid Fraud Control Unit at the Florida Attorney General's Office add significantly to the results of Medicaid fraud investigation in the state of Florida?***

In essence this demands a comparison of outcomes with and without the demonstration waiver, as illustrated in Figure 1, with exclusion or inclusion of the colored field named DMI.

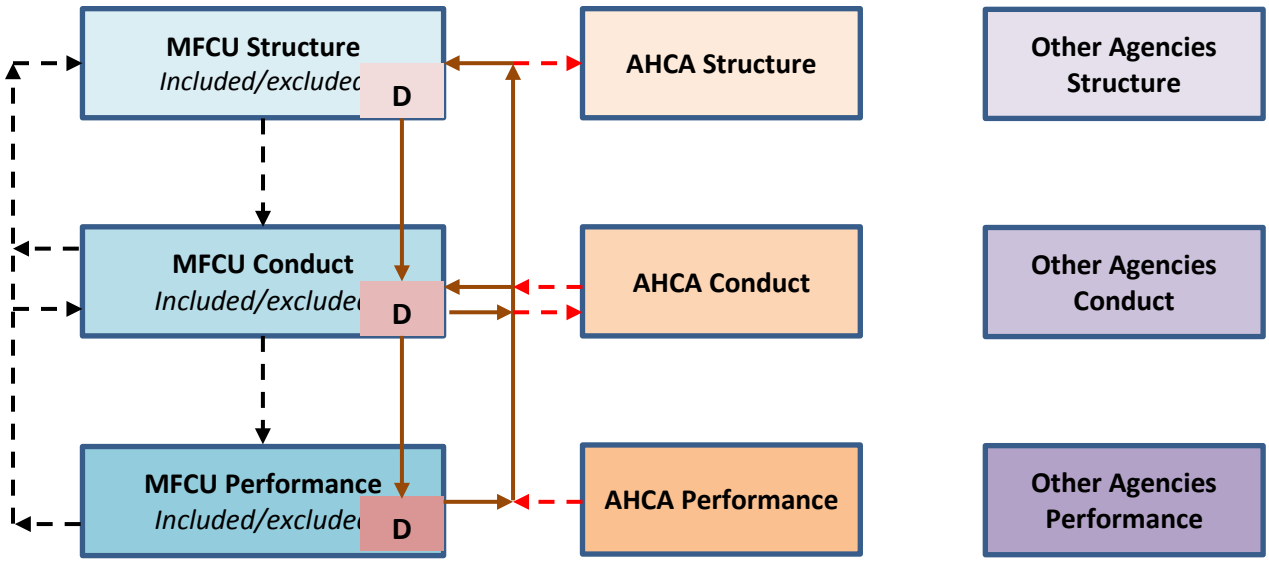


Figure 1: Structure-Conduct-Performance-Paradigm (SCPP) transposed on MFCU/DMI, AHCA and Other State and Federal Agencies.

The base framework used is the Structure-Conduct-Performance-Paradigm (SCPP) of Edward S. Mason.[7] According to this framework, an organization's performance depends on the conduct of its employees, which then depends on the structure. The reverse is also possible, e.g., once performance is determined or known, conduct and/or structure may change. In adding the Data Mining Initiative (DMI), based on the demonstration waiver and MOU, all levels will change inclusive between MFCU and DMI, as well as MFCU/DMI and AHCA (relevant arrows shown). The demonstration waiver, the MOU, and in particular the biweekly referral meeting and monthly data mining meeting (added structure elements) to discuss, interpret and exchange information on data mining projects (addition to conduct), enhances productivity and efficiency of MFCU and AHCA's resources (added performance). (Note: the red dashed arrows indicate the AHCA contributions on the various levels, as far as they pertain to the added DMI). Other agencies are also depicted in Figure 1, given that other agencies are consulted as well, but

[7] The paradigm was originally developed by Edward S. Mason, Harvard, in the 1930's. Since then it has been developed by J.S. Bain and other market structuralists in the field of Industrial Organization. It has also found use in amongst others the study of Economic Systems, and in Management and Organization.

arrows are omitted since these effects fall outside the scope of this evaluation. Noteworthy amongst others is also the commitment by MFCU to have adequate trained personnel as per the MOU (likewise an added structure element). In order to provide analyses on both scenarios (excluding versus including DMI), time series are used from FFY 2007-08 through FFY 2012-13, thus beginning a couple of years prior to the date that the demonstration waiver was granted.

In section 2, some available statistics are presented, relevant to the fraud investigation activities of the MFCU, including statistics of recent data mining activities. Preliminary results from interviews held with Key Informants on data mining are the subject of section 3. Section 4 covers the overall preliminary evaluation, and a preliminary interim conclusion is presented in Section 5.

2. Data Mining Activities Statistics

This section focuses on descriptive statistics based on data requests submitted to the Florida Attorney General's Office. It will cover more general statistics, as well as specific statistics on the data mining activities within the Medicaid Fraud Control Unit (MFCU). The purpose of presenting both types of statistics is to perceive the data mining activities in the proper relative context of the MFCU (as per Figure 1), as well as to present possible variables for the Data Mining Initiative (DMI) evaluation in section 4. This section will cover input variables (section 2.1), output variables (section 2.2), and outcome variables (section 2.3). Section 3 will cover the data mining process in further detail, based on interviews with key personnel and data mining analysts (akin to throughput variables). Figure 2 may help in perceiving the various variable categories in the proper setting. Given the variables, comparing input and output provides a measure of efficiency, while comparing input with outcome provides a measure of effectiveness. The presentation of data will be in accordance with the FFY, October 1st through September 30th.

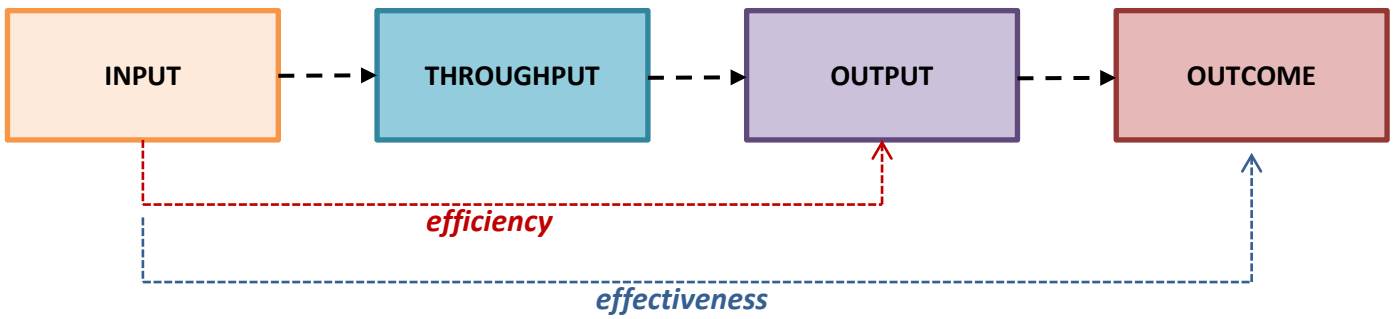


Figure 2: Input – Throughput – Output – Outcome Model.

2.1 Input: Budget, FTEs, and Training.

According to the requirements of federal statutes and regulations concerning the Federal Financial Participation (FFP), 75 percent of funding for the MFCU is provided by means of federal grants, and 25 percent are matching funds out of the State of Florida's General Revenue Fund and program income. Figure 3 depicts the MFCU budgets, inclusive of the FFP grants and the state matching funds, for FFY 2006-07 through FFY 2012-13. In addition, the MFCU funds provided through the FFP data mining grant (DMG) with matching state funds are included for FFYs 2010-11 through 2012-13.

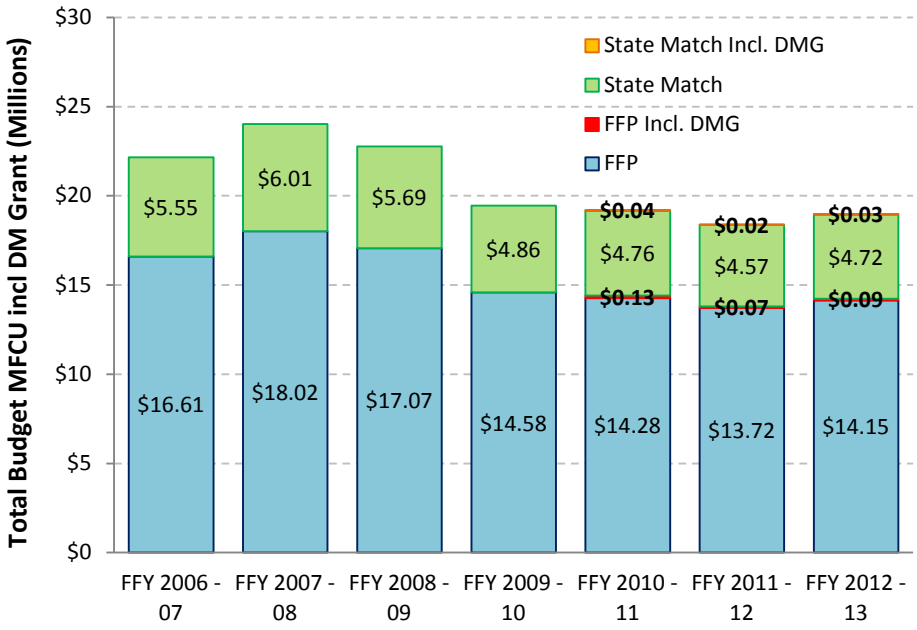


Figure 3: MFCU Budget, MFCU Grant and Data Mining Grant (Federal and State Matching Funds), FFY 2006-07 through FFY 2012-13.

As can be evidenced from Figure 3, the average total MFCU budget is approximately $20.5 million, with $15.4 million coming from the MFCU Grant and $ 5.1 million from State of Florida

matching funds. In focusing on the latter three years depicted, both FFY 2010-11 and FFY 2011-12 saw marginal budget declines, relative to the previous fiscal year budgets, while the budget for FFY 2012-13 came with a marginal increase. The added Data Mining Grants (both Federal Funding Participation (FFP) funds and Florida state matching funds), since FFY 2010-11, have had little impact on the budget and development thereof as mentioned, this given the relatively small contributions to the overall budget. The Data Mining Grant (DMG) therefore adds less than one percent (or approximately 0.7%) to the total MFCU budget. Figure 3a depicts the data mining budgets; including both FFP grant and Florida state matching funds, for FFY 2010-11 through FFY 2012-13.

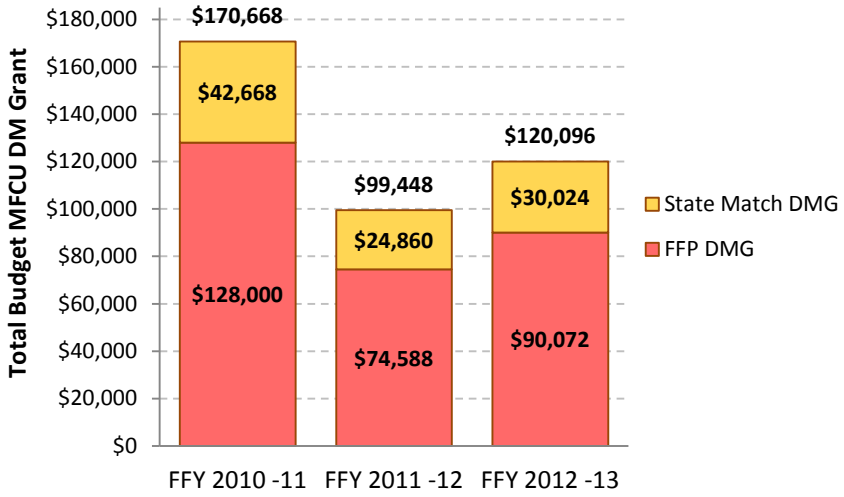


Figure 3a: MFCU Data Mining Initiative (DMI) Budget, Federal Data Mining Grant and Florida State Matching Funds, FFY 2010-11 through FFY 2012-13.

The lion's share, or 53.6 percent, of the FFY 2010-11 data mining budget was appropriated to "Equipment." The other two fiscal year budgets, namely FFY 2011-12 and FFY 2012-13, appropriated on average 51.3 percent of the respective budgets to "Salaries and Benefits." Although budgets are indicative for potential means of input, it is the actual allocation or expenditures that are relevant as a more direct input variable, and thus for the evaluation at hand. Both Figures 4 and 4a depict the differences between the budgets and expenditures, for

MFCU and DMI respectively, with the data on budgets from Figures 3 and 3a as a backdrop for comparative purposes. Both Figure 4 and 4a show that actual expenditures are less than the respective budgets. Data for the FFY 2012-13 is year-to-date (YTD).

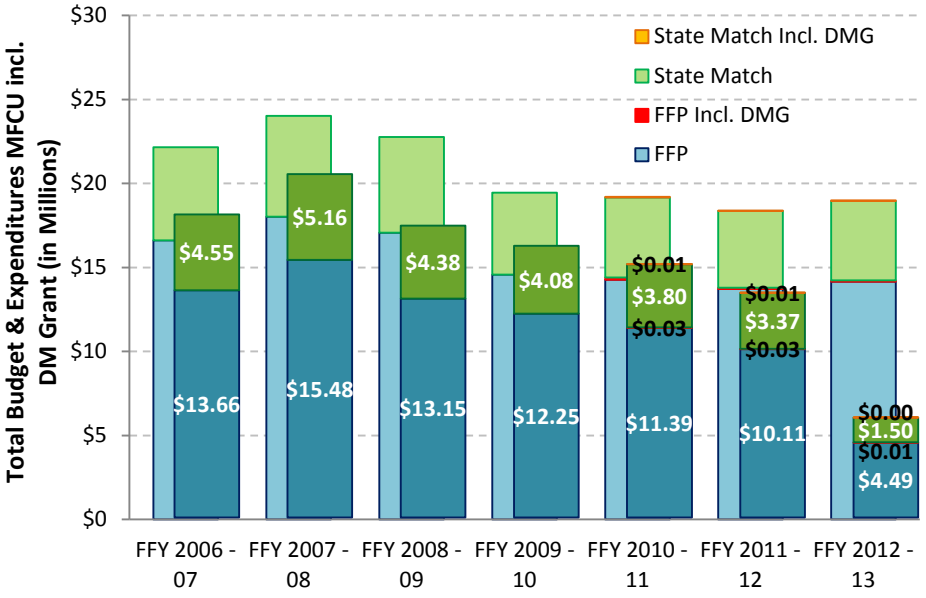


Figure 4: MFCU Budget and Expenditures, MFCU Grant and Data Mining Grant, FFY 2006-07 through FFY 2012-13 (YTD).

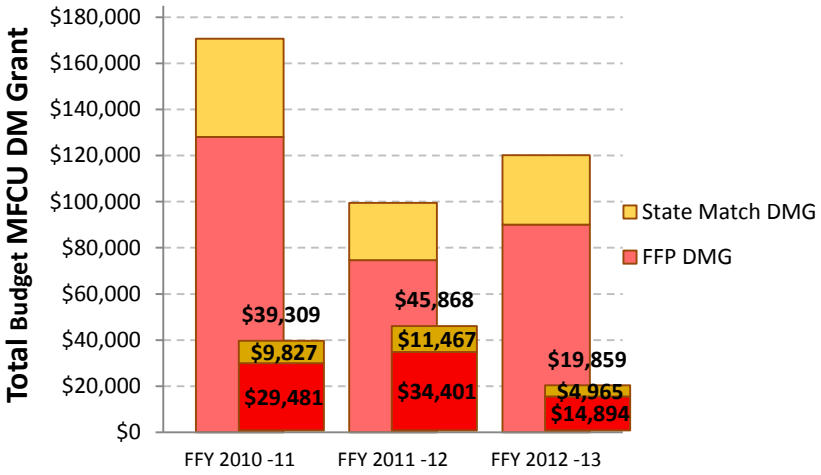


Figure 4a: MFCU Data Mining Initiative Budget and Expenditures, Federal Grant and Florida State Matching Funds, FFY 2010-11 through FFY 2012-13 (YTD).

Total expenditures by MFCU, on average, are approximately 80 percent of the fiscal year budgets, with a low of 73.5 percent for FFY 2011-12. For the DMI budget in particular (Figure 4a), expenditures come out at approximately 23.0 percent and 46.1 percent, for the two fiscal years available on the Data Mining Initiative (DMI). The lower level of expenditures is in part due to unfunded positions within MFCU.[8] As indicated, the specific expenditure data on both MFCU and DMI will be used as an input variable for the evaluation in section 4, albeit with some further corrections to be applied (e.g., to account for both time allocated for training and positions on reserve).

Table 1 provides some data on full-time equivalent (FTE) employment, both by type and by FFY 2006-07 through FFY 2012-13. The top row presents the total FTEs budgeted, while the second through fifth row provide a further breakdown by type of employment. The subsequent four rows give a breakdown and the total of FTE employment on reserve respectively, leading to a sub-total of FTEs applied or used by MFCU. Subsequently, the data mining analysts FTEs are added, from FFY 2010-11 onwards, resulting in total FTEs applied. Table 1a provides a further regional breakdown of data mining analysts by Florida MFCU region.

[8] Other reasons are as of yet unknown.

Table 1: MFCU Full Time Equivalent (FTE) Employment incl. Data Mining Analysts, Budgeted versus Applied, FFY 2006-07 through FFY 2012-13.

		FFY 2006-07	*FFY 2007-08*	*FFY 2008-09*	*FFY 2009-10*	*FFY 2010-11*	*FFY 2011-12*	*FFY 2012-13 (YTD)*
Total FTEs Budgeted		**232**	**232**	**232**	**217**	**214**	**210**	**210**
	Attorneys	26	26	26	27	27	27	27
	Investigators	131	131	106	101	100	97	97
	Auditors	7	7	7	7	10	10	10
	Support Staff	68	68	63	52	52	53	53
Reserve	Attorney			1	0	0	0	0
Reserve	Investigators	0	0	24	24	19	19	19
Reserve	Support Staff	0	0	5	6	6	4	4
				-30	-30	-25	-23	-23
Subtotal FTEs Applied		**232**	**232**	**202**	**187**	**189**	**187**	**187**
Data Mining Analysts Assigned FTE's (Tasks)						**0.45**	**0.75**	**0.75**
TOTAL FTEs Applied		**232**	**232**	**202**	**187**	**189.45**	**187.75**	**187.75**

Table 1a: MFCU Full Time Equivalent (FTE) Data Mining Analysts and Approximate Hours Devoted to Data Mining, per MFCU Region, FFY 2006-07 through FFY 2012-13.

DATA MINING GRANT					
		Region / Hours devoted to DMI			
	DMI Analysts FTEs	Northern Hours (%FTE)	Central Hours (%)	Southern Hours (%)	Total Hours (%)[9]
FY 2010-11	**0.45**	270 (15)	270 (15)	270 (15)	810 (45)
FY 2011-12	**0.75**	450 (25)	450 (25)	450 (25)	1.350 (75)
FY 2012-13	**0.75**	450 (25)	450 (25)	450 (25)	1.350 (75)

It is noted that the assigned data mining analysts FTEs (or better assigned data mining tasks) is quite small with respect to the overall MFCU employment, adding on average approximately 0.34 percent to the total formation. For evaluation purposes it is relevant to exclude the reserve FTE positions from input. In addition, on the data mining analysts' FTEs, it must be noted that two of the three original data mining analysts with the MFCU left the office in the course of FFY 2011-12.[10] The positions were filled by existing employees who were "brought up

[9] Calculus based on 1.800 hours per FTE.
[10] Exact timeframes are presently unknown, and thus its impact on applied FTEs/hours is still to be determined.

to speed" in a relative short timeframe. In principle, the input variable of data mining analysts FTEs needs to be adjusted for this timely impediment for further evaluation purposes in section 4. However, it was conveyed that little to no time was lost with the transition, and a qualitative judgment on difference in expertise and experience of data mining analysts with the specific data mining could not be made. Therefore no further adjustments are made, but the data presented needs to be valued in light of the transitions mentioned pro memory.

During the FFY 2011-12, the Medicaid Fraud Control Unit staff attended a total of 4,437.25 hours of training, while in FFY 2010-11 4,798.75 hours of training were attended. Given that there were 187 full-time employees (FTEs) assigned to the MFCU in FFY 2011-12, and 189 in FFY 2010-11, this means an average of approximately 23.6 and 25.3 hours in training per year respectively. Data mining analysts in particular attended 653.25 hours, 189 hours, and 66 hours (YTD) of training, during the federal fiscal years FFY 2010-11 through FFY 2012-13 respectively. Given that it doesn't make sense to divide the hours of training by FTEs, division by person delivers an average of 217.75 hours, 63 hours, and 22 hours (YTD) respectively for the data mining analysts.[11]

The focus of the MFCU data mining analysts' training in FFY 2010-11 was primarily on criminal analytics to increase the synergy between data mining activities and the fraud-oriented work context of the Florida Attorney General's Office, e.g., some 480 hours (or 73.5% of total training hours) were allocated towards "Florida Law Enforcement Analyst Training (FLEAT)." The main training batch of training hours was allocated towards Decision Support System (DSS) support contractor training (46 hours or 7.0%), followed by an Intelligence Officer Course (40 hours or 6.1%). In addition, seminars and webinars were attended. Main training providers were the Florida Department of Law Enforcement (FDLE), with 495 hours (or 75.8% of total training

[11] In taking approximately 1.794 hours per year for a full FTE, as per the Bureau of Labor Statistics, this comes out at 0.1214 FTE, 0.0351 FTE and 0.0123 FTE (YTD) per the fiscal years in consideration. Data retrieved from http://www.bls.gov/opub/mlr/2009/05/art1full.pdf

hours), and the AHCA, with 71 hours (or 11.3% of total training hours). Table 2 shows the top six course titles in training hours allocated in FFY 2011-12, and FFY 2012-13 (YTD) respectively. As seen from the table, the current scope of training is more diverse as compared to the first year of training, with its main emphasis of training on legal practices.

Table 2: Top Six Course Titles in Time Allocation for Training of MFCU Data Mining Analysts, FFY 2011-12 and FFY 2012-13 (YTD).

FFY 2011-12	percentage	hours
Financial Records Examination and Analysis - FREA	16.9%	32
Criminal Interview and Interrogations	12.7%	24
Tools of the Trade-Building Elder Financial Exploitation Cases	12.7%	24
Elder Abuse Training Program	8.5%	16
Certified Law Enforcement Analyst Training Seminar	8.5%	16
Courtroom Testimony	8.5%	16
	67.7%	**128**
FFY 2012-13 (YTD)		
Interactions between Medicaid Fraud Control Units and Program Integrity Units Symposium	36.4%	24
Cyber-Investigation 105 - Basic Cell Phone Investigations	24.2%	16
Exploring Interactive and Visual Data Mining	9.1%	6
Hemisphere Project	4.5%	3
State Medicaid Management Information System Long Term Care Training	4.5%	3
What Investigators & Analysts Need to Know about Facebook & Online Social Media: Awareness & Education Introductory Webinar	4.5%	3
	83.3%	**55**

2.2 Output: Complaints, Opened New Cases, Cases Investigated, and Disposition of Cases

Measures of outcome include numbers of "MFCU-opened" new cases, cases investigated, and cases closed. Complaints serve as the basis for most investigations done by the MFCU. During FFY 2011-12, the MFCU received a total of 1,317 complaints of which 292 (22.2%) were opened as operational cases. For FFY 2010-11 the MFCU opened a total of 354 new cases out of 1,661 complaints (or 21.3%). Data is depicted in Figure 5.

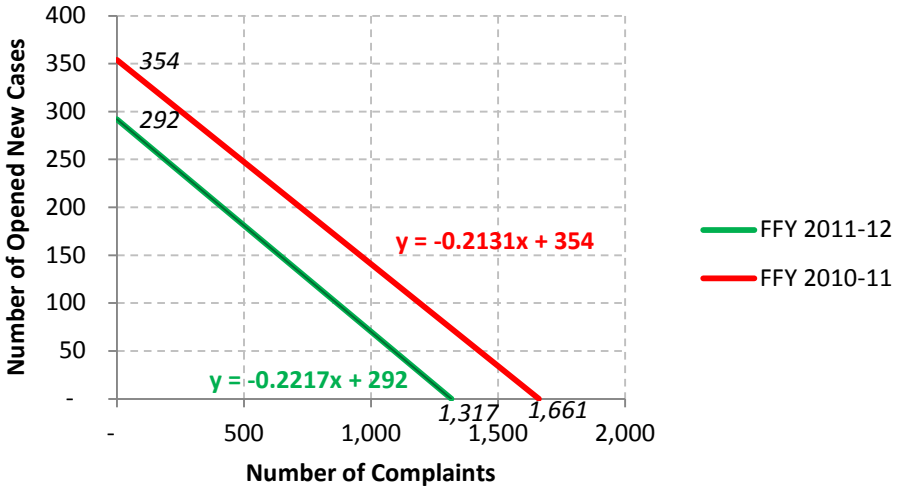


Figure 5: "MFCU-Opened" New Cases out of all Complaints, FFY 2010-11 through FFY 2011-12.

From Figure 5 it can be observed that the year to year opened new cases incidence ratio of opened new cases on complaints rose slightly from 0.2131 to 0.2217.

Table 3 provides data on the number of fraud complaints received by the MFCU. Average annual number of fraud complaints received by MFCU is 718 complaints (excluding FFY 2012-13).

Table 3: The Number of all Fraud Complaints Received by the MFCU, FFY 2006-07 through FFY 20120-13 (YTD).

Federal Fiscal Year	Number of Fraud Complaints Received
FFY 2006-07	498
FFY 2007-08	581
FFY 2008-09	510
FFY 2009-10	1171
FFY 2010-11	842
FFY 2011-12	707
FFY 2012-13 (YTD)	431

Table 4 on the next page gives an overview of the number of fraud complaints received by the MFCU, broken down by source, for the FFY 2010-11 through FFY 20120-13 (YTD). As can be evidenced from the Table 4, the number of complaints received by the source MFCU Data Mining Initiative is 27, 16, and 9 (or 3.2%, 2.3% and 2.1%) respectively for the three FFYs mentioned. Table 4a provides a selection of the same data i.e., the top eight sources of fraud complaints, with the MFCU Data Mining Initiative ranking as eight largest source, this based on relative averages for the three years FFY 2010-11 through FFY 2012-13 (YTD).

Table 4a: The Top Eight Sources by Number of all Fraud Complaints Received by the MFCU, Broken Down by Source, FFY 2010-11 through FFY 2012-13 (YTD).

	FFY 2010-11	FFY 2011-12	FFY 2012-13 (YTD)	Relative total by source FFY2010-11 through FFY 2012-13
Citizen	301	198	91	**29.8%**
Qui Tam	127	80	66	**13.8%**
Medicaid Recipient	50	108	95	**12.8%**
Family Member	22	82	69	**8.7%**
Employee	29	58	22	**5.5%**
AHCA - Medicaid Program Integrity	61	30	14	**5.3%**
Medicaid Provider	28	21	18	**3.4%**
MFCU Data Mining Initiative	**27**	**16**	**9**	**2.6%**
Total Number of Complaints	**842**	**707**	**431**	100%

Table 4: The Number of all Fraud Complaints Received by the MFCU, Broken Down by Source, FFY 2010-11 through FFY 2012-13 (YTD).

	FFY 2010-11	FFY 2011-12	FFY 2012-13 (YTD)		FFY 2010-11	FFY 2011-12	FFY 2012-13 (YTD)
AHCA - ALF Enforcement Unit			1	Family Member	22	82	69
AHCA - District Office	7	1		FBI - Federal Bureau of	4		3
AHCA - Fraud Prevention & Compliance Unit (FPCU)	2	8		FDLE - Florida Dept of Law	2		
AHCA - Health Quality Assurance	13	8	6	Government Employee	2	1	
AHCA - Medicaid Program Integrity	61	30	14	HHS - Health & Human Services	5	4	
AHCA - Office of Inspector General	3	3		HHS - OIG Health & Human Services	11	7	6
AHCA - Other Units	3	1		HMO - Investigative Unit	11	5	1
AHCA - Third Party Liability	1		1	Insurance Company	2	1	
Anonymous	13			Joint Task Force	4	1	
APD - Agency for Persons with Disabilities	20	10	6	Law Enforcement Agency	7	7	1
APS - Adult Protective Services	17	5	3	Medicaid Provider	28	21	18
Citizen	301	198	91	Medicaid Recipient	50	108	95
CMS - Center for Medicare & Medicaid Services	2			MFCU - Other than Florida	4	3	2
Confidential Informant	6	3		MFCU - Statewide Intel Team	2		
Consumer Protection Agency	1			**MFCU Data Mining Initiative**	**27**	**16**	**9**
Contractor for Center for Medicare & Medicaid	7	8	2	NAAG - National Association of	1	1	
County Health Department		1		NAMFCU - National Association of			
DEA - U.S. Drug Enforcement Agency			1	Operation Spot Check		1	
Dept of Children & Families - Inspector General	1			OSWP - Office of Statewide		1	
Dept of Children & Families - Other than APS	1	4		Press Report	2	4	
Dept of Elder Affairs	1			Qui Tam	127	80	66
DOH - Dept of Health	1	2		Social Security Administration (SSA)	1	20	
DOH - Medical Quality Assurance	1	2		Spinoff Case	31		13
DOJ - Dept of Justice	3			State Agency - Other	2		
DPAF Public Assistance Fraud		1		State Attorney's Office (SAO)	1		
Elected Official	2			U.S. Attorney's Office (USAO)		1	
Employee	29	58	22	Veteran Affairs			1
Transport				**Total Number of Complaints**	**842**	**707**	**431**

Table 5 shows the top five sources of fraud complaints received by the MFCU by provider, FFY 2010-11 through FFY 2012-13 (YTD).

Table 5: Top Five Number of all Fraud Complaints Received by the MFCU, Ranked by Provider, FFY 2010-11 through FFY 2012-13 (YTD).

FY 2010-11		
TOTAL	**842**	**Cumulative percentages of top 1 - 5**
Physician (MD) - 25	153	18%
Home and Community Based Service - 67	111	31%
Pharmaceutical Manufacturer	92	42%
Pharmacy - 20	64	50%
None	43	55%
FFY 2011-12		
TOTAL	**707**	
Physician (MD) - 25	123	17%
Home and Community Based Service - 67	99	31%
Pharmacy - 20	64	40%
None	48	47%
Dentist - 35	46	54%
FFY 2012-13 (YTD)		
TOTAL	**431**	
Physician (MD) - 25	62	14%
Dentist - 35	39	23%
General Hospital - 01	34	31%
Home and Community Based Service - 67	34	39%
Pharmacy - 20	34	47%

From the Table 5 it can be taken that the category Physicians (MD) ranks first in the three years depicted. Next, both Home and Community Based Services, and Pharmacy, show up in the top five of the three years. The last column of the Table 5 provides cumulative percentages on the top sources represented, showing that the top five providers represents a cumulative 55 percent, 54 percent and 47 percent of the total number of all fraud complaints received. Table 6 shows the sources of fraud complaints by provider type, where the source was MFCU Data Mining Initiative (DMI).

Table 6: Number of Fraud Complaints Received, by Provider Type, where the Source was Data Mining Initiative, FFY 2010-11 through FFY 2012-13 (YTD).

Federal Fiscal Year	Number
FFY 2010-11	**27**
Physician (DO) - 26	4
Physician (MD) - 25	21
Therapist (PT, OT, ST, RT) - 83	2
FFY 2011-12	**16**
Home and Community Based Service - 67	12
Physician (MD) - 25	1
Therapist (PT, OT, ST, RT) - 83	3
FFY 2012-13 (YTD)	**9**
Dentist - 35	9

For the Data Mining Initiative (DMI), the largest provider category of fraud complaints was Physician (MD). The second largest provider category is Home and Community Based Service, while Dentist is the third largest category for the DMI.

Of the complaints mentioned only a subset will result in case status (for processes see Section 3). Table 3 provides information on MFCU cases investigated and opened new cases by source (sources defined per agency/category), fiscal years FFY 2006-07 through FFY 2012-13 (YTD).

Table 7: MFCU Cases Investigated, Cases Opened, and the Source of the Cases, FFY 2006-07 through FFY 2012-13 (YTD).

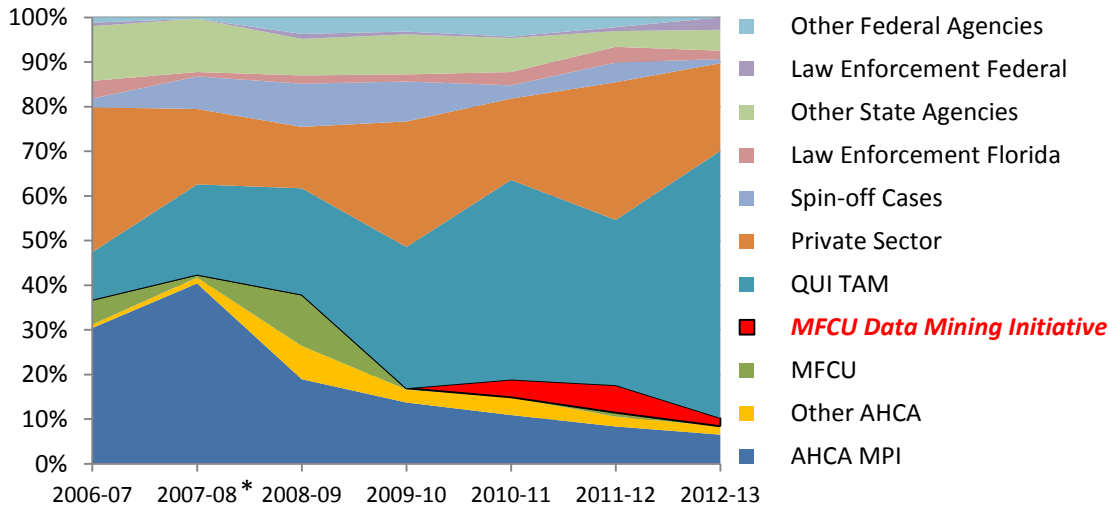
	Federal Fiscal Years						
	FFY 2006-07	FFY 2007-08	FFY 2008-09	FFY 2009-10	FFY 2010-11	FFY 2011-12	FFY 2012-13*
Cases: Investigated**	927	922	927	906	930	872	
Cases: Opened New During FFY	253	302	269	313	354	292	107
Cases: Sources of New Opened Cases (sources defined by agency):							
AHCA - Medicaid Program Integrity	77	122	51	43	33	19	7
Other AHCA	2	4	20	9	12	5	2
MFCU	14	2	31	1		2	0
MFCU Data Mining Initiative					12	14	2
Qui Tam	27	61	64	99	135	84	64
Private Sector	82	51	37	88	55	70	21
Spin-off Cases	5	22	26	28	9	10	1
Law Enforcement Florida	10	3	5	5	9	8	2
Other State Agencies	31	36	22	28	23	8	5
Law Enforcement Federal	2		3	2	1	2	3
Other Federal Agencies	3	1	10	10	13	5	0

*YTD **Caseload is a snapshot of the number of cases on the last day of the Federal Fiscal Year.

As per Table 7, the average number of cases investigated is approximately 914 cases per year (excluding FFY 2012-13). Similarly, on average approximately 278 new cases are opened during a fiscal year. The major sources of new opened cases are *qui tam*[12] and Private Sector sources (e.g., citizens, employees, providers, recipients, contractors, media), at a relative average of approximately 30.1 percent and 22.8 percent respectively. The third largest source of opened

[12] *Qui tam* is a lawsuit brought by a private citizen (popularly called a "whistle blower") against a person or company who is believed to have violated the law in the performance of a contract with the government or in violation of a government regulation, when there is a statute which provides for a penalty for such violations. Qui tam suits are brought for "the government as well as the plaintiff." In a *qui tam* action the plaintiff (the person bringing the suit) will be entitled to a percentage of the recovery of the penalty (which may include large amounts for breach of contract) as a reward for exposing the wrongdoing and recovering funds for the government. Sometimes the federal or state government will intervene and become a party to the suit in order to guarantee success and be part of any negotiations and conduct of the case. This type of action is generally based on significant violations which involve fraudulent or criminal acts, and not technical violations and/or errors. http://dictionary.law.com/default.aspx?selected=1709

new cases is the AHCA with a relative average of approximately 22.9 percent; 19.9 and 3.0 percent for AHCA-Medicaid Program Integrity and Other AHCA respectively. MFCU comes in at a relative average of approximately 2.8 percent of opened new cases, with DMI at 4.4 percent (based on FFY 2010-11 through FFY 2012-13 YTD only). DMI added 4.1 percent to the sub-total of opened new cases in FFY 2010-11, 6.6 percent of opened new cases in FFY 2011-12, and 1.9 percent of opened new cases in FFY 2012-13 (YTD). On the source or action initiating data mining, complaints are by far the prime driver of new activities, while pending (criminal) cases are next. The same data as Table 7, on opened new cases by MFCU per source, is depicted in Figure 6 in relative terms (FFY 2012-13 YTD).



** In FFY 2007-08, biweekly briefings began between AHCA MPI and MFCU with an emphasis on the quality of referrals being made to MFCU.*

Figure 6: Relative Shares of Opened New Cases by Source, FFY 2006-2007 through FFY 2012-13 (YTD).

Table 8 provides a further breakdown on opened new cases by region; DMI opened new cases versus all other cases sources, for FFY 2010-11 through FFY 2012-13 (YTD).

Table 8: Opened New Cases by Region; DMI and Other Sources, FFY 2010-11 through FFY 2012-13 (YTD).

		FFY 2010-2011		FFY 2011-2012		FFY 2012-2013 (YTD)		FFY 2010-2011	FFY 2011-2012	FFY 2012-2013 (YTD)
CCEB		135	135	89	89	64	64			
Central	DMI opened	7		6		2		58.3%	42.9%	
	Other opened	54		47		17		34.8%	37.9%	41.5%
			61		53		19			
Northern	DMI opened	3		7		-		25.0%	50.0%	
	Other opened	56		42		12		36.1%	33.9%	29.3%
			59		49		12			
Southern	DMI opened	2		1		-		16.7%	7.1%	
	Other opened	45		35		12		29.0%	28.2%	29.3%
			47		36		12			
	DMI opened	12		14		2				
	Other opened	290		213		105				
	DMI/Other*	4.1%		6.6%		1.9%				
Total			**302**		**227**		**107**			

* DMI/Other opened = 12/ (302-12), 14/ (227-14), and 2/ (107-2)

The middle columns of Table 4 show the number of DMI-attributed opened new cases (sienna colored rows) and all other sources opened new cases (blue colored rows), adding to the total in the last row of the table. As can be observed in Table 4, Complex Civil Enforcement Bureau (CCEB) is the largest source for opened new cases, with a relative average of 45.3 percent of total MFCU opened new cases for FFY 2010-11 through FFY 2012-13 (YTD). The spread of opened new cases over the MFCU regions is quite even, with Central Florida at a relative average of 20.9 percent, North Florida at 18.9 percent, and South Florida at 14.9 percent. The last three columns provide the relative shares of opened new cases per region, excluding the CCEB opened new cases (e.g., 7 / 12 = 58.3%; 54 / (290-135) = 34.8%; et cetera). The relative shares indicated in red, show that the regional DMIs added relatively more out of the DMI-opened new cases to the region, than other sources did out of all other sources. The variable "opened new cases" will be used for evaluation purposes in section 5.

Table 5 provides a list of the top five Medicaid Provider types for Medicaid fraud ranked from most to least frequency of fraud.

Table 9: Top Five of Medicaid Fraud Cases by Provider Type, FFY 2010-11 through FFY 2012-13.

Fraud Cases Opened by Provider Type		
FFY 2010-11	FFY 2011-12	FFY 2012-13 (YTD)
• Pharmaceutical Manufacturers • Home & Community Based Services • Physicians (MD) • Pharmacy • General Hospital / Therapist	• Home & Community Based Services • Pharmaceutical Manufacturers • Physicians (MD) • Pharmacy • Medical Equipment Manufacturer	• Pharmaceutical Manufacturers • Pharmacy • General Hospital / Physicians (MD) / Medical Equipment Manufacturer • Home & Community Based Services • Independent Lab

From Table 9, it can be observed that Pharmaceutical Manufacturers, Home and Community Based Services, and Physicians (MD), lead in number of opened new fraud cases according to rank numbers. Of cases attributed to the DMI, Physicians (MD), Physicians (DO), Therapists, Home and Community-Based Services, and Dentists are the main categories of opened cases by provider type. Given that cases by provider type can only be measured in frequency or rank number, this variable will not be used for further evaluation in section 4.

Table 10 gives an overview of the disposition of MFCU cases closed, as well as the subset of cases closed attributed to the Data Mining Initiative (DMI), FFY 2010-11 through FFY 2012-13 (YTD). Shade formatting in the table is provided to make a visual distinction between lower counts (blue fields), higher counts (brown and orange fields), and median counts (white fields).

Table 10: Disposition of MFCU Cases Closed and Subset of Cases Closed Attributed to the Data Mining Initiative, FFY 2010-11 through FFY 2012-13 (YTD).

	MFCU			of which: DMI		
Cases: disposition of Closed Cases	FFY 2010-11	FFY 2011-12	FFY 2012-13 (YTD)	FFY 2010-11	FFY 2011-12	FFY 2012-13 (YTD)
Administrative Closure	32	2	4			
Administrative Referral	65	55	23	1	2	
Assistance to Other Agencies		1	8		1	
Case Dismissed	22	11	6			
Case Remanded	3					
Civil Intervention Declined	5	1	1			
Civil Judgment	2	2				
Civil Settlement	45	14	17			
Consolidated	16	3	4			
Conviction	24	9	5			
Defendant Deceased			1			
Defendant filed Bankruptcy	1					
Lack of evidence	28	23	12	4	3	
Nolle Prosequi	2					
Plea Agreement	7	10	9			
Pretrial Intervention	3	2	3			
Probation			1			
Prosecution declined		6				
Resolved with Intervention	1	2				
Unfounded	18	25	9		1	
Voluntary Dismissal	11	21	24			
Grand Total Closed Cases	**285**	**187**	**127**	**5**	**7**	**0**

As can be observed from the table, only a subset of MFCU cases lead to settlement, conviction, or plea agreement. Administrative referral is 22.8 percent and 29.4 percent of MFCU cases, for FFY-2010-11 and FFY 2011-12 respectively. For the DMI these percentages are 20.0 percent and 28.6 percent respectively. Of MFCU cases, 9.8 percent and 12.3 percent are closed due to lack of evidence, in FFY-2010-11 and FFY 2011-12, respectively. Similarly, of the DMI cases 80.0 percent and 42.9 percent are closed for the same reasons. Given that the disposition of cases

closed can only be measured in frequency or rank number, this variable will not be used for further evaluation in section 4.

2.3 Outcomes: Monies Recovered

A longer term perspective on outcomes of activities by the MFCU, in terms of total amount of the monies recovered, is presented in Figure 7. Two regression lines are depicted next to the total amounts recovered, an exponential and a straight line regression. In using the exponential regression (with R^2 = 0.8484), it can be derived that the average growth in recoveries has been 26.1 percent annually. The straight regression line (with R^2=0.7823) is drawn to provide credence to the perception that recoveries might not grow as exponentially going forward.

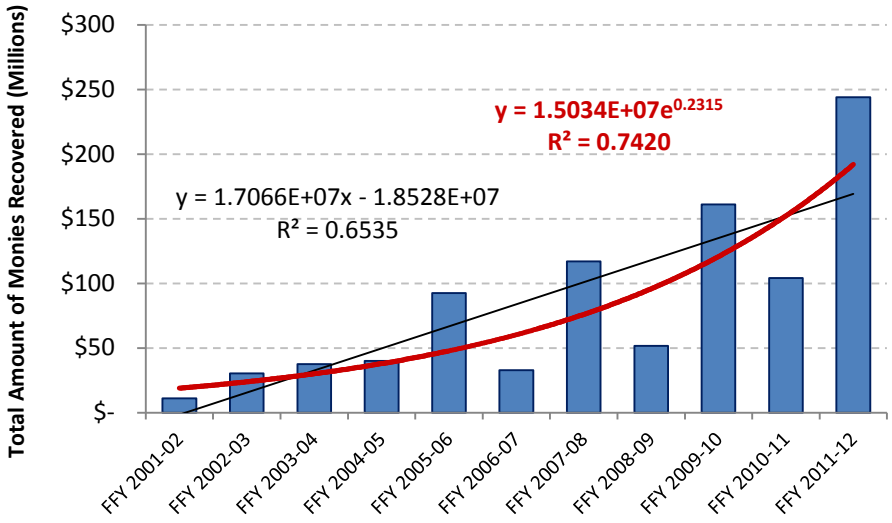


Figure 7: Total Amount of Monies Recovered by MFCU, FFY 2001-02 through FFY 2011-12.

Figure 8 compares the number of cases investigated to the total amount of monies recovered by MFCU.[13]

[13] Figure 8 and relevant narrative still based on State fiscal year (SFY).

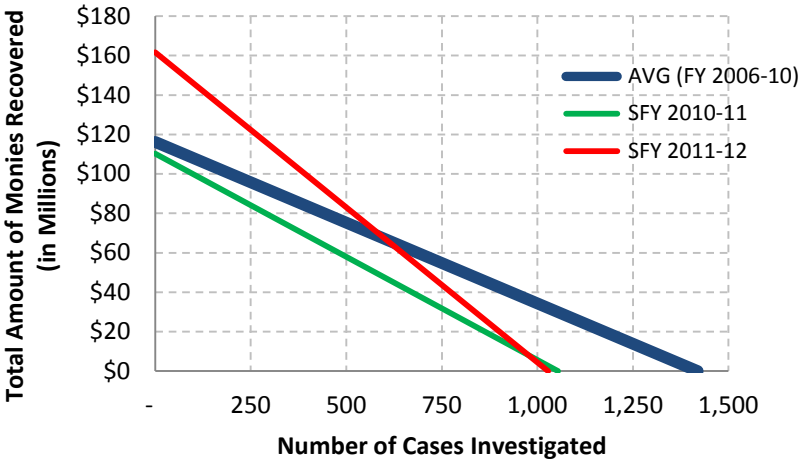


Figure 8: Number of Cases Investigated Relative to the Total Amount of Monies Recovered in Millions, Average SFY 2006-10, SFY 2010-11 and SFY 2011-12.

The bold line represents the average ratio of Total Amount of Monies Recovered versus Investigated Cases for the SFYs 2006-07 through SFY 2009-10 (1,419 cases versus $116.2 million in total recoveries). In SFY 2010-2011, MFCU recovered a total of $110.3 million on 1,054 investigated cases. With almost equal recoveries, and approximately a quarter less in number of investigated cases, this means a higher average recovery ratio on investigated cases. Similarly for SFY 2011-12, the number of cases investigated is 1,028, while the total sum of recoveries came in at $161.7 million. With an almost equal number of investigated cases the total monies recovered rose by slightly over 46 percent (approximately 46.6%). In short, the steeper the angle, the higher the ratio of monies recovered over cases investigated.

In the Florida state fiscal year (SFY) 2011-12, the total amount for civil recoveries, which include civil settlements arising from *qui tam* cases brought under Florida's False Claims Act, was $145,374,604.[14] The total for criminal recoveries based upon Medicaid fraud cases was $14,020,038.65. The total amount of monies recovered by the MFCU in SFY 2011-12 was

[14] Figure 9 and relevant narrative still based on State fiscal year (SFY).

$161,667,067.25. In addition, the MFCU's recoveries generated $22,720,363.51 through penalties imposed and $37,431.82 in interest that was deposited into the State of Florida's General Revenue Fund. The total amount of monies recovered by the MFCU for SFY 2010-11 was $110,276,959. The amount for civil recoveries by the MFCU in SFY 2010-11 was $107,079,438, and the amount for criminal recoveries based upon Medicaid fraud was $3,197,521. The Unit recoveries generated $16,414,495 through penalties imposed and $467,243 in interest deposited. Figure 9 depicts the same total amount of monies recovered per SFY relative or next to the respective input or budgetary means, total Federal Financial Participation (FFP) and Florida General Revenue Funds or Program Income. In taking the values from Figure 9, the year-to-year rise in total recoveries constitutes approximately 47 percent.

Figure 9: Total Amounts of Monies Recovered and Total FFP + Florida, SFY 2007-08 through SFY 2011-12.

In SFY 2011-12, for every FFP and General Revenue dollar spent, the MFCU generated approximately $5.54 through penalties and interest deposited into General Revenue.

To date, 12 cases attributed to the DMI have been brought to a close, and came with the following dispositions: administrative referral, assistance to other agencies, lack of evidence, or were unfounded (as per Table 6).

3. Data Mining Activities Key Informant Experiences – Preliminary Findings.

This section is mainly based on interviews with key personnel at both the Medicaid Fraud Control Unit (MFCU) and the Florida Agency for Health Care Administration (AHCA), as well as sitting in on inter-agency meetings. The purpose is to derive a clear perception of the means-end decision or process chain, from pre data mining activities within the MFCU to the perceived inter-agency communications and cooperation between MFCU and AHCA. The inter-agency communications are based on the biweekly meeting, as well as the monthly data mining meeting (added structure elements based on the memorandum of understanding (MOU)).

It needs to be mentioned that even before commencement of the Data Mining Initiative (DMI), senior management teams for both AHCA and MFCU, as well as the Department of Health (DOH), met on a monthly basis to discuss major issues, strategies, joint projects and other relevant matters. The objective in describing the activities/inter-agency activities is to find aspects relevant to the evaluation (Structure-Conduct-Performance Paradigm), and possibly potential recommendations to improve upon the data mining process within the MFCU. The following narrative will focus on the MFCU data mining analysts first, MFCU staff second, and on conduct and interactions between the organizations MFCU and AHCA third.

A questionnaire was developed with a list of semi-structured questions for interview purposes to get a clear perception of the process. For the data mining analysts, the semi-structured interview questions were categorized in such a way as to shed light on the following aspects of data mining:

1) Research team,

2) Procedures and protocols,

3) Queries, algorithms and models,

4) Validation,

5) Documentation or filing of practices, and

6) Other more general questions.

MFCU: Data Mining

1) RESEARCH TEAM

The data mining analysts in the workforce at the MFCU increased from 0.15 FTE to 0.25 FTE in each of the South, Central and North Florida offices, from Federal fiscal year (FFY) 2010-11 to FFY 2011-12. Before commencement of the Data Mining Initiative (DMI), as per October 2010, all three data mining analysts were power users with the Florida Decision Support System (DSS). The term "power user" is used to indicate the highest level of data mining analysts (based on adequate training), who have priority in data access and analyses. Since the three data mining analysts became part of the MFCU, they received 653.25 hours, 189 hours, and 66 hours (YTD) in training for each of the three years (FFY 2010-11 through FFY 2012-13) under consideration. As indicated, law enforcement criminal analyst training by the Florida Department of Law Enforcement (FDLE) constituted the major focus of training in FFY 2010-11. Subsequent training covered a variety of applied and practical issues (see Table 2).

Prior to October 2010, the research team at MFCU had access to the DSS databases (with billing and other information), but any data mining activity had to be either case specific or be based on an allegation or complaint. As with production, every subsequent project under research or investigation leads to added learning experiences by the data mining analysts (learning curve), e.g., raised understanding, new acquired perceptions, and gained insights. This learning leads not only to improved skills, but above all to a derived product or effect (spin-off). However,

added data mining activities, based on improved skill and "spin-off" (also described as "what if" questions), were not allowed if the activities did not meet the condition of research on pending and specific cases only. These potential added data mining activities (also referred to as "phishing") could only be communicated with and referred to AHCA, with the result that outcomes of data mining were received from AHCA with a time-lag only. In consequence, the learning curve gain which usually results in higher productivity was interfered with structurally. This may have resulted in not only delayed learning (and loss of learning), but also to less added productivity by the data mining analysts at MFCU, which is a loss of potential fraud or abuse cases. In addition some information may be lost in communication and the subsequent data mining activity (it is noted that files are communicated or exchanged, not methodology or results of data queries).

The present procedure under the waiver, with checks by AHCA on possible duplication (as taken from the interviews with key personnel) works quite efficiently. The direct personal communication on proposals at the biweekly meetings, (to discuss interpret and exchange information and perceptions on data mining projects, adding verbal information to potential projects) gives ample time to learn and understand the objectives and expectations of each agency. In addition to the bi-weekly meetings, analysts made reference to the monthly meeting between MFCU and AHCA on data mining, with a likewise and candid exchange on data mining issues. Increased synergies are mentioned in interviews with key personnel at both MFCU and AHCA. Both meetings seem to be highly valued, from both agency's perspectives. In short, the present procedure works fairly smoothly, and fairly efficiently. The biweekly meetings lead to an exchange of information on what everybody is doing.

2) PROCEDURES AND PROTOCOLS

The initial "trigger" for data mining analyses can be an idea (learning experience), a concept, or a person/provider, and can either be based on a complaint or pending case. Proposed or suggested data mining activities or projects by MFCU are relayed to AHCA at the biweekly meetings. AHCA distributes the suggested project to other relevant AHCA staff and vendors, and replies to MFCU usually within the timeframe of one week. This relay is instituted to check with the different agencies on whether there is an issue of duplication of data mining activities. Potential projects denied to date: Eleven out of 71 potential cases have been denied to date.

On each potential project, two checks are performed; the first is on the promise of outcome, and if promising, the data mining needs are put in queue with a tracking number and log. The second check is on whether a person/provider is already under investigation. Concerning the latter, data mining activities may add information to an open case, or potentially designate an offender as a repeat offender. Once a data mining activity by MFCU is commenced, a project file is set up. Each project is entered into the Data Mining Initiative (DMI) Tracking Log, whether approved or denied by AHCA, both for tracking and historical purposes. This DMI Tracking Log currently is in Microsoft Excel format.

3) QUERIES, ALGORITHMS AND MODELS

Different data mining techniques are used on the DSS Databases, utilizing tools such as amongst others Microsoft Excel, Access Pivot, and Phi2 (mainly by AHCA). Programmed algorithms (beyond Microsoft Excel functions) are not used and are perceived to be the prerogative of the support contractor (Hewlett Packard). On the question of whether the data mining activities could best be described by: (A) "statistics, neighborhood and clustering," or (B) "trees, networks and rules," univocally the answer was both. Outlier analysis is generally perceived as a first data mining analysts task only, and usually is a data summarization/aggregation tool, while data mining thrives on detail. Further diving into more detail or particular data was considered

necessary to look for patterns, e.g., trending, spikes, and out of the ordinary claims. Even with scale issues (large versus small providers) and/or scope issues (specialists versus general providers), data mining activities can be quite focused on provider type, type of service, specialty of medical provider, timeframe, and/or geographic locations.

4) VALIDATION

Once a data query is run and data is retrieved, the results are documented in a Data Mining Analyst Report (DMAR) with a DMAR track number. The translation by the data mining analyst, from data mining output to a report being written with recommendations, is the first step in deriving information from the data. This translation determines the further cause of the data mining analyses project in terms of justification, and for deciding whether to drop it, refer it to AHCA, or move it to the next level as a potential law enforcement issue. The latter usually will lead to further communications with the data mining analyst, on which there may be repeated rounds of data mining activities. Given the data mining analyst reports available in queue, validation is typically done by MFCU staff based on different perceptions, inclusive of legal and medical expertise. Similarly, justification is sought in filed complaints,[15] which may precede a determination to case level. It is noted that it takes time to prepare and process a legal dossier, even long after the data mining activities are done. Any subsequent involvement of law enforcement leads to a full-blown case. However, if deemed truly administrative at any stage, the project or case will be closed by the MFCU and referred to AHCA.

5) DOCUMENTATION OR FILING OF PRACTICES

All analysts' activities are accounted for in Data Mining Analyst Reports (with DMAR number, and/or subsequent OAG file-number), filed in the system, and put in the Tracking Log. Queries and models are saved and can be run again either at will or at selected regular intervals. All

[15] A complaint is an allegation that a person or provider may have committed an offense that may constitute a violation of state or Federal law.

data mining activities are reported and filed, name or case specific (as per legal practice) as key for potential later use.

An investigative report on Data Mining Initiative (DMI) comprises the following sub-tabs:

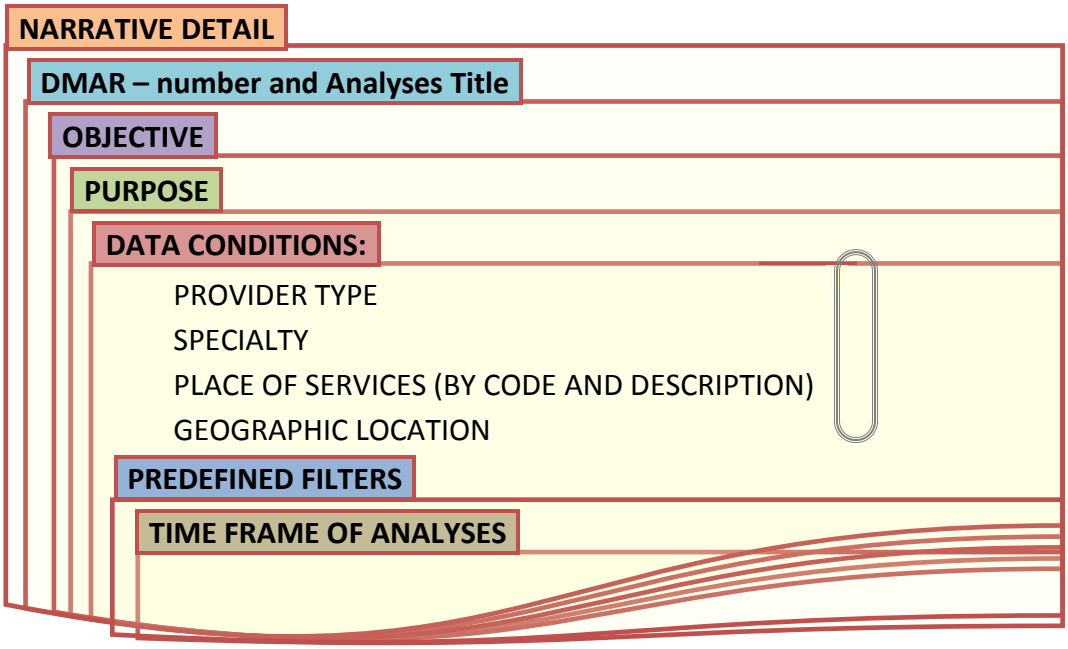


Figure 10: Various Tabs of an Investigative Data Mining Activities Report.

Once a project becomes a case, the DMAR report is combined with further investigative and legal documentation, and filed in the computer-based case management system with an OAG tracking number. This system comes with various sub tabs as well; namely, summary, contacts, investigation, status, legal status, supplemental information, attachments, evidence, and statistics.

6) OTHER MORE GENERAL QUESTIONS

Links Analyses Software was mentioned as a data mining tools/software that may be helpful for the Attorney General Office's data mining activities, and which is currently not available or in

use. Links Analyses is a VisuaLinks® - Link Analysis Software, a platform-independent, graphical analysis tool used to discover patterns, trends, associations and hidden networks in any number and type of data sources.[16] In substantiating the need, reference was made to 1) the higher volume of activities with an added number of projects, 2) more complete and robust package for tracking (instead of the presently used MS Excel), and 3) the need to generate forms and letters for AHCA, and potentially other agencies, all with increasing accessibility for future purposes.

Overall, the perception was that with the DMI, data needs were more readily met (as compared to the prior "data on request only" structure with the AHCA), that response time on "what if" data needs decreased dramatically, and supportive data mining in pending investigations readily added information to cases. The position of AHCA is fully recognized, understood and highly respected with its responsibilities and specialist expertise. The objective is to work on fraud and abuse, while the MFCU's focus is on criminal activity.

MPI/MFCU Bi-weekly Meeting and DMAR Meeting

In experiencing the MPI/MFCU Bi-weekly meeting, referral discussions went swift and with clear assignment. Under the label of topics and other discussions, various issues were exchanged in a manner of not only adding and exchanging information from various fields of expertise, whether it was medical, Medicaid protocols, legal perspectives, experience or otherwise, but quickly building a comprehensive perception on each issue. The direct accommodative and supportive communications lead to quick and increased insights for everybody present. Any other form of communication, even e-mail, between the organizations

[16] Visual Analytics Incorporated (VAI) is a leading provider of information sharing and visual data mining products. VisuaLinks presents data graphically, uncovering underlying relationships and patterns. VisuaLinks addresses the entire analytical process – from access and integration to presentation and reporting – providing a single and complete solution to a broad range of data analysis needs. For more information see: http://www.visualanalytics.com/products/visualinks/index.cfm

to achieve similar results would have taken quite a longer timeframe. In principle the meetings are an added learning curve experience, increasing expertise on handling cases and issues, and thus increasing the efficiency of means allocated by both organizations. The DMAR meeting was different in the sense that it was not the singular cases, but common denominators that were exchanged. These common denominators were the different data mining options, but also methodologies, opening new avenues and opportunities for further data mining activities.

AHCA:

The main focus of the interviews held with AHCA staff was on the interaction with MFCU. It was revealed that the director of Data Detection left the office October 2012, and the position was not filled as of this date (May 2013). Operational communication between MPI and MFCU however continues, especially with regards to the scheduled biweekly MPI/MFCU meeting addressing the project requests by MFCU (concerning the issue of potential duplications), and the monthly data mining meetings.

MPI does extensive Medicaid research on providers, practices, claims and billing, as well as payments based on its administrative, legislative, market and medical expertise, and drawing on its team of specialists. On data mining, MPI uses the DSS and has direct access via desktop/server. In addition to Microsoft Excel and IDS, Active Data Base software is used (which is deemed better than Access Pivot). Results of data mining activities by the "Detection Group" are forwarded to the "Case Management Group." This group decides on further handing; i.e., dropping the project, additional records request, processing, or referring to the MFCU if deemed potentially fraudulent. On referral to the MFCU the files are shared (not the data mining queries). All projects, inclusive of MFCU referred cases, are tracked by number.

Incoming data mining project requests from the MFCU, prior to the waiver, were put in queue since limited resources are allocated to the most promising projects first. In addition, the so-

called “power users” have priority in data mining, and thus overrule access by others. Under the waiver, incoming proposals are checked both internally and externally with other agencies for possible duplication of data mining activities. MFCU is notified usually within the timeframe of one week on its data mining requests.

The waiver allowing MFCU to data mine is seen as an additional opportunity to face abuse and fraud in a more involved manner. Data mining by MFCU is not seen as competition but as a partnership with mutual rewards in terms of getting resolve on abuse and fraud. The communication between the offices adds to the information stream and increases insights on potential issues. The data mining activities by MFCU are perceived not as full-fledged investigations, but more as auxiliary investigations with the main intent to support activities within the Florida Attorney General’s Office. With the waiver it has become possible to communicate between the two organizations on a different level, and exchange information without duplication of AHCA expertise into the office of the Florida Attorney General or, vice versa, law enforcement expertise into AHCA. Overall the waiver was considered to add to the mutual working relationship between AHCA and the MFCU. System improvements could be made by setting up a Sharepoint portal (as mentioned in the interviews).

Based both on the waiver and the MOU signed between the Office of the Attorney General and the Florida AHCA , new structures and procedures had to be put in place which define and determine the position, and to a certain extent, the conduct of the MFCU. In addition, budgetary requirements had to be met. It is observed here, that as a consequence of the waiver, the learning curve experience has improved (shorter response time form AHCA), and in addition, a new inter-agency learning experience is created by interpretation and information exchange at a high specialist data mining analyst level, between the two organizations at hand.

4. Data Mining Activities Preliminary Evaluation

On the evaluation of the Data Mining Initiative (DMI) at the Medicaid Fraud Control Unit (MFCU) at the Florida Attorney's Office, the question is whether or not the data mining waiver, as a demonstration project, added significantly to the results of Medicaid fraud investigation in the state of Florida. As per Figure 1 it was discussed that DMI can neither be seen apart or isolated from the activities within the MFCU, nor from the inter-agency activities with the Agency for Health Care Administration (AHCA). Second, there are some limited variables to provide some static measure of efficiency and effectiveness (as per Figure 2).

Figure 11 shows a recap of some of the key output data points, or achievements, from Section 2, providing both the numbers on the axes (with the right-hand side of the horizontal axis having two scales, one on complaints, and one on cases ending in settlement, conviction, or plea agreement) and perceptions on the ratio's; 1) complaints/fraud complaints, 2) fraud complaints/opened new cases, 3) opened new cases/cases disposed, and 4) cases disposed/cases ending in settlement, conviction, or plea agreement, for the FFY 2010-11 and FFY 2011-12 consecutively (similar to Figures 5 and 8). For instance in FFY 2010-11, reading the figure counter clockwise, a total of 1,661 complaints were received (first scale on the right hand side of the horizontal axis), some 842 fraud complaints were dealt with (top vertical axis), 354 new cases were opened (left hand side of the horizontal axis), and some 285 cases were disposed (bottom part of the vertical axis). Finally, some 76 cases were brought to a settlement, conviction, or plea agreement (second scale on the right hand side of the horizontal axis). Consequently, the ratios 1 through 4 (depicted by the slopes) are: 842/1,661 = 0.5069, 354/842 = 0.4204, 285/354 = 0.8051 and 76/285 = 0.2067. Similarly, for FFY 2011-12, a total of 1,317 complaints were processed, some 707 fraud complaints were handled, 292 new cases were opened, and some 187 cases were brought to a close. In addition, some 33 cases ended in a

settlement, conviction, or plea agreement. The FFY 2011-12 ratio's therefore are: 707/1,317 = 0.5368, 292/707 = 0.4130, 187/292 = 0.6404 and 33/187 = 0.1765.

The ratio's are all below one, since it should be clear that complaints outnumber cases, and not all cases come with an arrest, or a positive outcome in terms of monies recovered. It must be noted that Figure 11 depicts parallel FFY data only, and not successive (or causal) results from complaint to disposition, or tracking of complaints over the years, from complaint to disposition. Figure 11 maps the year-to-year activities of the MFCU on all fronts (data and ratios); activities on which time and other resources are allocated, to review, refer, work with the investigative team, et cetera.

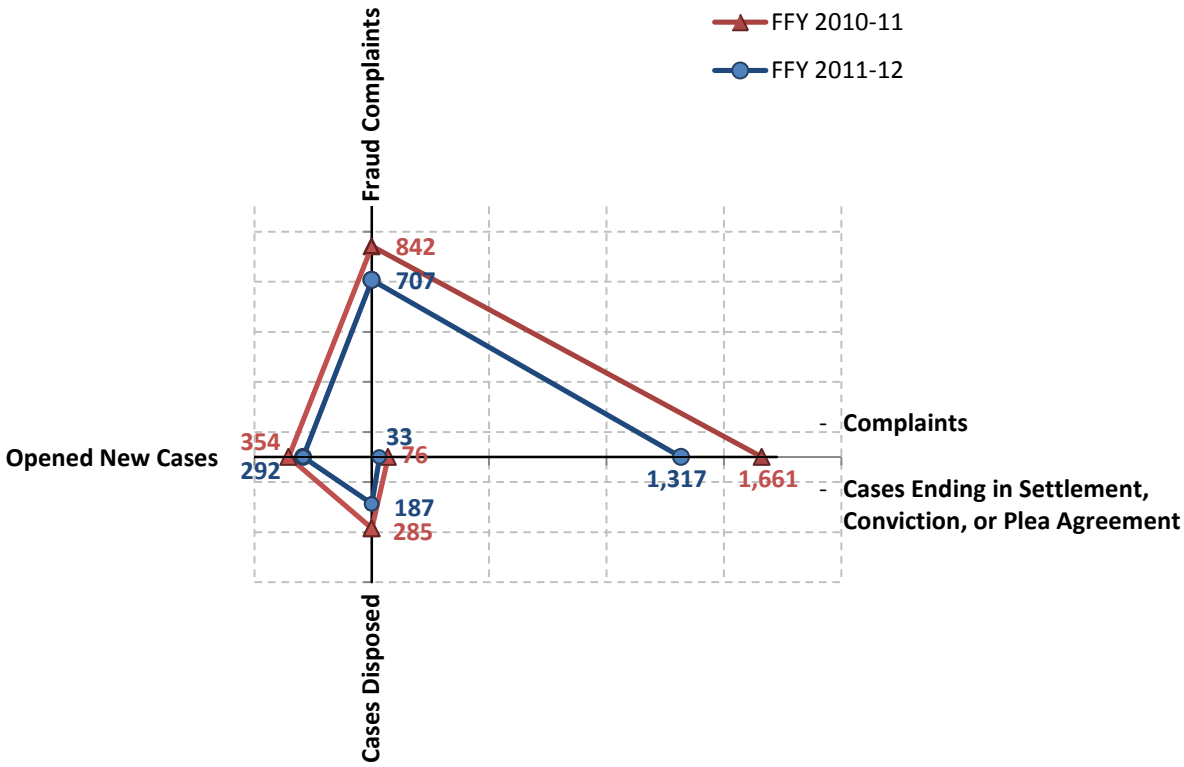


Figure 11: Number of Complaints, Opened New Cases, Disposition of Cases, and Cases Ending in Settlement, Conviction, or Plea Agreement, MFCU, FFY 2010-11 and FFY 2011-12.

Given marginal or small differences, slightly higher ratios on Fraud Complaints, Opened New Cases, Cases Disposed and Settlement, Conviction Plea Agreements for FFY 2010-11, and a slightly higher ratio on Fraud Complaints/Complaints for FFY 2011-12, the two maps show quite similar FFY activity patterns.

A similar set-up for the MFCU Data Mining Initiative (DMI) is given in Figure 12.

Figure 12: Number of Complaints, Opened New Cases, Disposition of Cases, and Cases Ending in Settlement, Conviction, or Plea Agreement, Attributed to DMI, FFY 2010-11 and FFY 2011-12.

From the Figure 12 it can be taken that the incidence ratio of Opened New Cases over Fraud Complaints changed quite dramatically from FFY 2010-11 to FFY 2011-12 (12/27 = 0.4444 and 14/16 = 0.875 respectively) and is high in comparison to the same incidence ratio of Figure 11 (354/842 = 0.4204 and 292/707 = 0.4130 for FFY 2010-11 and FFY 2011-12 respectively).

In using a time series analyses, in particular a multi regression analyses with the DMI as an added variable, it may be possible to derive some preliminary insights (given the short timeframe) from a more dynamic perspective. This brings a hypothetical element in the evaluation, which is to value and compare output/outcomes under different scenarios; namely, with and without the DMI under the waiver. For evaluation purposes, the perception is taken that the waiver provides an opportunity (e.g., data mining as a working tool) for the Attorney General's Office to increase the efficiency of labor input. DMI efforts (FFY 2006-07 through FFY 2011-13 YTD) are captured, by making the number of opened new cases dependent on the total budget and DMI adjusted FTEs (increased efficiency of labor with the DMI tool), according to the following format:[17]

$$\text{Opened New Cases} = a * (\text{FFP} + \text{FL.GR})^{\alpha} * (\text{FTE}^{\beta} * \text{DMI}^{\gamma})$$

in which:

FFP + FL.GR = Federal Financial Participation (FFP) and Florida General Revenue/Program Income means, expenditures only (in real prices of 2012),[18]

FTE = Effective employment in FTEs,[19]

DMI = Data Mining Initiative adjustment margin on FTEs.[20]

[17] For some preliminary analyses on the equation see the appendix.

[18] Annual budget data adjusted with Price Indexes for Gross Domestic Product according to Table 1.1.4. Price Indexes for Gross Domestic Product, Bureau of Economic Analyses, http://www.bea.gov, date retrieved April 15, 2013.

[19] FTEs are adjusted for time allocated to training. For the MFCU, excluding Data Analysts, 22 hours or 0.0123 FTE are assumed from each FTE for FFY 2006-07 through FFY 2008-09 and 20 hours or 0.0111 FTE for each FTE for the fiscal years FFY 2009-10 onwards. For the data analysts 0.1214 FTE, 0.0351 FTE and 0.0123 FTE per analyst per fiscal years FFY 2010-11 though FFY 2012-13 is take for training purposes. In addition for FFY 2012-13 only half a year is assumed.

[20] For years without DMI a dummy variable of 1 is used (i.e., no impact). For years with DMI an adjustment margin is used. The margin for FFY 2010-11 is taken at 0.9603 (or 1/(1+12/(302-12))), for FFY 2011-12 at 0.9383 (or 1/(1+14/(227-14))), and for FFY 2012-13 at 0.9813 (or 1/(1+2/(107-2))) as per DMI assigned opened new cases.

Given the equation, the number of opened new cases (or output) stands in direct relation to the expenditures and FTEs adjusted by a DMI factor (or input factors). The equation allows the DMI to be analyzed in conjunction with the FTEs, with DMI as an added tool to increase the efficiency of labor input. Therefore, the equation brings to the fore the essence of the evaluation issue and allows for sensitivity analyses i.e., varying one variable while leaving the others constant (ceteris paribus). A multiple regression analyses on the data points FFY 2006-07 through FFY 2012-13 (YTD) results in an expected number of new cases according to the format:[21]

$$\text{Opened New Cases} = 0.4696 * (\text{FFP} + \text{FL. GR})^{08432} * (\text{FTE}^{-1.4465} * \text{DMI}^{0.7567})$$

t-Stat	-1.0606	3.5578	-1.9808	0.4645
P-value	0.3667	0.0379	0.1419	0.6739

With $R^2 = 0.9657$ and Adj. $R^2 = 0.9314$

Figure 13 displays the actual versus the expected number of new cases, based on the multiple regression equation calculated, for the fiscal years FFY 2006-07 through FFY 2012-2013 (YTD).

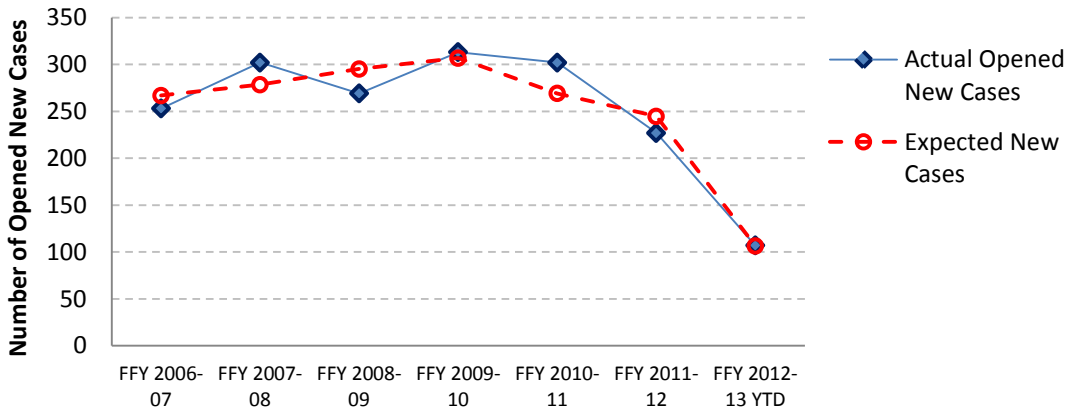


Figure 13: Actual versus Expected Number of Opened New Cases MFCU, FFY 2006-07 through FFY 2012-13 (YTD).

[21] Regression calculus done is preliminary, given that data for FFY 2012-13 is YTD, and FTEs for FFY 2012-13 is taken at half the budgeted value.

Given the equation, it is possible to conduct a sensitivity analyses, varying one variable while keeping other variables constant, measuring the impact on the output or opened new cases. Figure 14 provides the results of a sensitivity analyses done with available data.

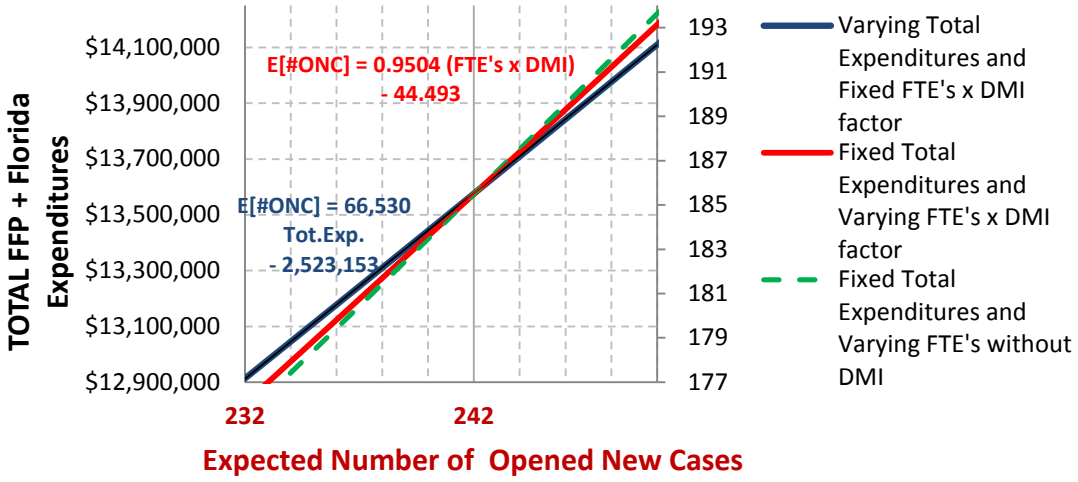


Figure 14: Sensitivity Analyses of Average Budget and Full Time Equivalent Employment on Expected Number of Cases.

The intersection in Figure 14 represents the present (FFY 2011-12) position with a total allocated expenditure at $13,580,769 (left hand scale), 185.56 applied FTEs (corrections from 187.75 FTEs due to training (right hand scale)), leading to a regression estimated number of 242 opened new cases (as per the realized 227 opened new cases in FFY 2011-12). From this present point, first the variable Total Expenditures (FFP + FL.GR) is changed within the range of plus to minus five percent, under ceteris paribus condition (i.e., leaving other variables constant), with results presented by the series "Varying Total Expenditures and Fixed FTEs x DMI factor." As can be taken from Figure 14, the positive slope of the total expenditure line means that an increase in expenditures (left hand vertical axes), will raise the output in terms of number of opened new cases (horizontal axes). More precisely, a one percent increase in

expenditures will raise the number of opened new cases by approximately 0.8419 percent (inelastic). Secondly, the DMI factor is varied (third variable in the equation, range ibid, under ceteris paribus). Since the DMI factor is taken in combination with the FTEs, this makes for the series "Fixed Total Expenditures and Varying FTEs x DMI factor." Similarly the positive slope of the line means that an increase in DMI will raise the output in terms of number of opened new cases (horizontal axes). In particular, a one percent increase in the DMI factor will raise the number of opened new cases by approximately 0.8050 percent (inelastic).
The dashed line represents opened new cases due to changes in FTEs without DMI (as per the situation of FFY 2011-12).[22] The dashed line has an elasticity of approximately 0.75409. Graphically, the lower the slope of the line, the higher the impact of a change is on the number of opened new cases. In short, the right personnel is more important than expenditures, and good personnel combined with the right tools such as DMI only improves upon the output, this by approximately 6.64 percent (0.8419/0.7409).

No relation is found between any measure of input and cases investigated. Data on cases investigated are a snapshot in time only, as per the close of the fiscal year.

No relation is found between any measure of input and cases closed.

No relation is found between any measure of input and monies retrieved. The explanation is that no measure for recoupment is attributable as of yet to DMI, since the program is still in its infancy. In addition, and more general, it may be that the order or outcome of investigations doesn't come with a similar range of values or monies retrieved, since the fraudulent entities and the order of fraudulent activities may differ in size and scope.

[22] Dashed line is obtained by transposing the situation as per FFY 2011-12, and varying the DMI margin in the regression equation from the value of 1, i.e., for years without DMI.

5. Interim Conclusion

This report presents an evaluation of the MEDS-AD waiver: Data Mining Activities, contingent on the waiver CFR 1007.19. With respect to the evaluation, the question is:

> ***Did the Data Mining Initiative (DMI) at the Medicaid Fraud Control Unit at the Florida Attorney General's Office add significantly to the results of Medicaid fraud investigation in the state of Florida?***

Given that the Data Mining Initiative (DMI) cannot be seen apart or isolated from the activities conducted within the Medicaid Fraud Control Unit (MFCU) of the Attorney General's Office, the framework used is the Structure-Conduct-Performance-Paradigm (SCPP), with DMI as an add-on to the MFCU. Various input, output and outcome variables available were looked at for properly representing the relative position of data mining activities. Descriptions were given on input variables: expenditures, FTEs, and training, from both MFCU and DMI. Output variables, especially cases investigated, opened new cases and closed cases, were looked into, and finally the outcomes in terms of monies recovered.

Static analyses showed a slight rise in the incidence ratio from 0.2028 to 0.2282 of opened new cases on number of complaints. The number of complaints received by the source MFCU Data Mining Initiative is on average 2.6 percent annually. Opened new cases attributed to the DMI showed an average of 4.4 percent of total opened new cases, over the three years of evaluation. The ratio of total amount of monies recovered over cases investigated showed a clear increase over the years (average FFYs 2006-10, FFY 2010-11, FFY 2011-12). Dynamic analyses indicates that expenditures are inelastic at 0.8419 with respect to opened new cases, while the DMI adjustment factor (adjusting FTEs for becoming more efficient) proved inelastic as well at 0.8050, this in terms of output or number of opened new cases. Therefore, the right

personnel is more important than expenditures, and good personnel combined with the right tools such as DMI only improves upon the output in terms of opened new cases. More specifically, the dynamic analyses show that DMI add approximately 6.6 percent of opened new cases, which is slightly higher than the static number of 4.4 percent mentioned.

No dynamic relation is found between any measure of input and cases investigated, cases closed, or monies retrieved.

A special concluding note must be made on the improved learning curve experience as a consequence of the waiver and MOU between the MFCU and AHCA. In addition, a new inter-agency learning experience is created by interpretation and information exchanged at the high specialist data mining analyst level.

Appendix: Production Function Used

A production function is taken to be:

$Y = f\ (K, L)$

where:

Y = total output or outcome produced in a year,

K = capital input; in this evaluation total expenditures,

L = labor input; effective FTEs per year.

Comparison will be made with:

$Y = f\ (K, L')$

in which in addition:

L' = adjusted or augmented labor input due to the DMI.

In particular a Cobb-Douglas production function is used in the format:

$Y = a\ K^{\alpha} * L^{\beta}$

in which in addition:

a = total factor productivity

α and β are the output elasticities of capital and labor, respectively. These values are constants determined by available technology.

For the purpose of this evaluation the production function is rewritten in the format:

$$\text{Opened New Cases} = \text{a} * (\text{FFP} + \text{FL. GR})^{\alpha} * (\text{FTE}^{\beta} * \text{DMI}^{\gamma})$$