RESEARCH REPORT MAY 2012

Evaluation of the Forensic DNA Unit Efficiency Improvement Program

By:

David Hayeslip

Sara Debus-Sherrill

Kelly A. Walsh

URBAN INSTITUTE
Justice Policy Center

URBAN INSTITUTE
Justice Policy Center

2100 M Street NW
Washington, DC 20037
www.urban.org

This project was supported by Contract No. GS-23F-8198H awarded by the National Institute of Justice. The National Institute of Justice is a component of the Office of Justice Programs, which also includes the Bureau of Justice Statistics, the Bureau of Justice Assistance, the Office of Juvenile Justice and Delinquency Prevention, and the Office for Victims of Crime. Points of view or opinions in this document are those of the authors and do not represent the official position or policies of the United States Department of Justice, the Urban Institute, its trustees, or its funders.

ACKNOWLEDGMENTS

The Evaluation of the Forensic DNA Unit Efficiency Improvement Program was funded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The authors are grateful to Katharine Browning, Mark Nelson, and NIJ for their guidance and assistance in facilitating the evaluation. We also wish to wholeheartedly thank each of the grantee sites for the generous time donated to this project's goals. Their commitment to this study was shown in numerous hours spent explaining and demonstrating grant projects, monthly phone conversations to update the research team on grant activities, and retrieval of data to objectively document the impacts of the program. At the Urban Institute we would like to thank Susan Narveson for her extensive participation in this evaluation as a forensic expert consultant and for reviewing drafts of this report, as well as Akiva Liberman and P. Mitchell Downey for lending their expertise when statistical guidance was needed. Finally, we credit Rayanne Hawkins and Shalyn Johnson for their administrative help in bringing this project to completion.

Contents

List of Figures

List of Tables

EXECUTIVE SUMMARY

The Urban Institute contracted with the National Institute of Justice (NIJ) to conduct an evaluation of its 2008 Forensic DNA Unit Efficiency Improvement Program. This executive summary describes the background of this demonstration program, the scope of the evaluation and its methodology, key implementation and outcome findings, as well as cross-site conclusions. The full evaluation report provides additional background about DNA and its use in criminal investigations, details about the research methodology and findings, as well as supplementary individual site documentation.

The NIJ Forensic Unit Efficiency Improvement Program

In May 2008, NIJ issued a competitive solicitation to provide funding to public crime laboratories for the implementation of new and innovative approaches designed to improve the efficiency of DNA evidence processing. This initiative was significantly different than past federal efforts, which primarily focused on providing funding to reduce DNA evidence processing backlogs through increased capacity. Capacity-building efforts generally focused on improving infrastructure, information management systems, operations, automation and improved evidence storage. In contrast, the Unit Efficiency Program was designed to focus directly upon laboratory evidence processing, including the identification of "bottlenecks" and the application of holistic approaches to increase efficiency rather than just increasing capacity.

Of the 13 public crime laboratories that applied to the program, six were funded: Allegheny County Medical Examiner's Office, Forensic Laboratory Division (PA); Harris County Medical Examiner's Office Forensic Biology Laboratory (TX); Kansas City Police Department Crime Laboratory (MO); Louisiana State Police Crime Laboratory (LA); San Francisco Police Department Criminalistics Laboratory (CA); and University of North Texas Center for Human Identification (TX). Each grantee committed to a 25 percent local match, and total project funding (federal and nonfederal) ranged from $120,000 to $1,365,956 per site.

A wide variety of activities was proposed in order to increase efficiency at these six crime labs. Examples include work flow process mapping, automation with new robotic applications, expert systems, new chemical procedures, laboratory information management system (LIMS) improvements, document management systems, and other software applications. Of the six original grantees, two ultimately withdrew from the program. Harris County, TX, decided not to participate shortly after receiving its grant award, and San Francisco, CA, withdrew during the early stages of implementation.

The Scope of the Evaluation

The Urban Institute's primary goal was to systematically evaluate the implementation and potential outcomes of the proposed innovative approaches to improve DNA crime laboratory efficiency. This was first accomplished through an assessment of implementation at each of the sites (Harris County was excluded since it withdrew, but San Francisco was included due to partial implementation). The research team collected and analyzed data from a wide variety of site-specific documents, interviews, and on-site observations. Outcomes were assessed in several different ways. Using site LIMS processing data, case and sample productivity (i.e., throughput and turnaround time) were measured. In addition, efficiency indices were created to examine changes in productivity as a function of resource units (personnel and budgetary) that might be associated with the novel grant activities. Productivity and efficiency were also assessed in the context of key implementation milestones. This was done through visually examining longitudinal trends in relation to implementation milestones, pre/post comparisons utilizing t- and Mann-Whitney U tests, and regression analytic techniques. Stage-to-stage changes were examined, where feasible, along with start-to-finish DNA processing.

It should be noted that this evaluation did not examine the effectiveness of individual interventions, nor their performance or validation testing in controlled experimental laboratory settings. Instead, this evaluation focused on observable effects of the NIJ Forensic DNA Unit Efficiency Program in the real-life settings and functions of operational laboratories.

Key Findings

There were significant implementation delays across the study sites. In fact, each of the sites had to request no-cost grant extensions from NIJ because of their implementation challenges. In addition, some of the delayed components did not become fully operational until after the evaluation was complete, which necessarily limited outcome measurement and assessment. However, it should not be surprising that such delays were encountered. The demonstration labs, as is the case for crime labs nationwide, have been facing exponentially increasing requests to process DNA samples. Not only has this been due to the increasing demands associated with the growing importance of DNA evidence in criminal casework, but also because of the added workload associated with arrestee and convicted offenders testing. Turnover of key project personnel was a challenge, particularly for one site which had its project director leave in the middle of implementation. Other factors beyond lab control, such as cumbersome and time-consuming procurement regulations and demands from external accreditation, also affected the timeliness of implementation. The unique nature of requirements for process and equipment validation and staff training also played a role.

Project management also varied across sites, particularly in terms of strategic planning and extent of collaboration. It appeared that laboratory-wide engagement was related to greater implementation success. On the other hand, "vertical" project management with a single leader did not appear to be as effective. For instance, one site's project director left, leaving a large knowledge void that temporarily stopped progress. Another site encountered problems implementing a fully validated process into casework, partially due to a lack of coordination with casework leaders during the beginning stages. In addition, data-driven approaches to the identification of processing bottlenecks and addressing changes based upon careful data analyses appeared most effective in creating solutions tailored to efficiency improvements.

It was also found that data entered into and maintained in crime lab LIMS were fraught with difficulties from both evaluation and lab management purposes. While useful perhaps from a day-to-day operational perspective, LIMS are not effective tools for performance monitoring and measurement due to a lack of electronically recorded key information, inconsistent data entry, and data field overwrites. In addition, linkages to agencies submitting DNA samples were often limited, making communication about case status more difficult. However, it is noted that very few crime laboratories provide direct LIMS access to these agencies. Significant improvements may be made in the quality of internal lab data to facilitate DNA processing performance monitoring, management and outcome research.

Productivity and efficiency outcome findings were mixed across the sites. There were considerable month-to-month and case-to-case variations in lab processing statistics both within and across the study sites. In addition, the non-linear nature of processing, which often can include stage reruns, illustrated the complexity of DNA processing work and how outcomes can vary drastically from case to case.

Nonetheless, there was clear evidence of significant increases in throughput and turnaround time in Louisiana. In addition, the analysis of efficiency indices revealed positive outcomes for this site. Throughput somewhat increased in Kansas City but proved not to be statistically significant; analyses of turnaround time showed mixed findings. For the other two remaining sites, the post-implementation period was too short to adequately assess outcomes statistically, although there was some initial evidence of potential success.

Conclusions

The findings of the Evaluation of the Forensic DNA Unit Efficiency Improvement Program suggest that there is some evidence in support of the hypothesis that DNA processing can be improved in novel and innovative ways above and beyond simply increasing capacity. Due to implementation challenges and methodological limitations, the findings may be best viewed as a conservative estimate of the short-term outcomes of the grant program. Regardless of measured outcomes, significant scientific contributions to the field were made through

participating labs attempting something innovative. Examples include demonstrating how organization-wide changes can be made, validating steps that can be taken to decreasing time-consuming steps in DNA processing, expanding the kinds of systems and chemistries that are acceptable as valid field practices, and making this information available publicly to other labs, among others.

The results of this evaluation also clearly show how important future research is for both the social science and physical science fields in this area. Understanding the interrelationships between capacity, productivity, and efficiency appears particularly important for policymakers and practitioners in order to make the most informed choices about how to address processing backlogs and bottlenecks in crime laboratories. Improving knowledge about how to address crime lab challenges should match the increase in the importance of forensic science and the demands placed upon it by the criminal justice system. As the criminal justice system continues to rely more heavily on the forensic sciences, laboratories will need to pursue new solutions to growing organizational demands. Understanding which of these solutions is most effective is important for both the forensic science and criminal justice fields, as well as the community at large.

Evaluation of the Forensic DNA Unit Efficiency Improvement Program

INTRODUCTION

The use of Deoxyribonucleic Acid (DNA) typing to aid criminal investigations has expanded significantly since routinely being applied to casework by the FBI in the late 1980s. It is widely accepted as an extremely accurate method of identification. Unfortunately, the rapid growth in the number of requests for DNA analysis has far exceeded the processing abilities of many forensic laboratories. As a result, DNA processing backlogs have been growing, a recent estimate of which was almost 100,000 requests in 2008 (Nelson 2010). This backlog estimate does not include evidence still in police possession and not yet submitted by investigators, which, in 2003, NIJ estimated to include 350,000 rape and homicide cases (NIJ 2003).

In an attempt to alleviate this growing DNA evidence processing backlog, the President's DNA Initiative was launched in 2003. Under this program, federal grant funds have been provided to crime labs throughout the country, most of which were designed to increase lab capacities through infrastructure, personnel, technology, and other resources. The FY2008 National Institute of Justice (NIJ) Forensic DNA Unit Efficiency Improvement Program was an attempt to address the backlog problem from a different perspective, focusing on efficiency gains rather than improvements in capacity. Under this demonstration program, six crime labs from around the country were selected to receive funding to acquire, validate, and implement innovative strategies to improve efficiency.

The Justice Policy Center of the Urban Institute (UI) was awarded a contract from NIJ to conduct an evaluation of the first year of the NIJ Forensic DNA Unit Efficiency Improvement program. The goal of this evaluation was to systematically evaluate both the implementation and outcomes of this program. While six crime labs were initially funded by NIJ, two of the sites subsequently withdrew from the program. The evaluation includes documentation of the implementation of this grant program at each of the funded sites. In addition, an assessment of how the program affected laboratory DNA request processing was also conducted at each of the four final sites. The researchers examined both the stage and overall laboratory *productivities*, as well as stage and overall laboratory *efficiencies* (productivity by resource units).[1]

In this final report we describe the novel efficiency strategies developed at each site, the implementation of program components overall and within each site, present results about

[1] For a more detailed discussion of the difference between *productivity* and *efficiency* as they are used in this report, please see section 1.5.

the program productivity and efficiency outcomes within each lab, examine what was learned across the sites, and provide recommendations for the implementation of future crime lab efficiency efforts and additional research.

The results of this evaluation have potentially far-reaching implications for DNA processing, particularly relative to increasing throughput through efficiency instead of just increased capacity. Innovative approaches to increasing processing efficiencies may hold great promise for improved criminal investigations and prosecutions in the future. In addition, this evaluation helps provide a foundation for future research, from both social science and physical science perspectives, into the use of crime laboratory innovations that might change how evidence is processed through the use of technology and other approaches to accommodate the continued future growth in the use of DNA evidence and the demands placed on crime labs nationwide.

1. BACKGROUND

1.1 DNA Evidence and Forensic Laboratory Processing

1.1.1 Deoxyribonucleic Acid (DNA)

Deoxyribonucleic Acid is the genetic material ultimately responsible for all inherited traits. Structurally, this chemical is built like a ladder, with sides made up of phosphate-sugar complexes and the rungs made of paired chemicals called bases. These bases are adenine, thymine, guanine, and cytosine, commonly abbreviated A, T, G, and C. On the "rungs" of the DNA ladder, A pairs with T and G pairs with C. The sequence of these bases holds the coded information necessary for cells to build proteins. These proteins ultimately perform the functions in the human body that result in physical and biological traits. In the cell, the DNA ladder is twisted and coiled into 46 separate chromosomes (23 chromosome pairs consisting of one chromosome of each pair inherited from the mother and the other chromosome from the father of the individual). This set of chromosomes is collectively referred to as the human genome.[2]

Throughout the human genome there are sections of DNA that are not known to code for any physical or biological traits and/or functions. Within these non-coding sections are patterns of bases (2–7 base pairs long) that are repeated numerous times (e.g., AATG.. AATG..AATG..AATG..) at specific locations (loci) on the DNA molecule. These sections are called short tandem repeats (STRs). They are the target of most forensic nuclear DNA processing. Laboratory procedures are designed to isolate the DNA molecule from the biological matrix, isolate and make copies of targeted STRs in the molecule, and separate

[2] For additional information about DNA structure and function, please visit http://www.dna.gov/basics/biology/.

these copies based on their number of repeats. The data produced are used to determine the number of times the short tandem sequence pattern is repeated at each locus.

The final product of forensic STR processing is the DNA profile. The profile is a series of numbers, where each one represents the number of repeated patterns of DNA at a particular location on the DNA molecule. The profiles, produced from forensic evidence, are compared to profiles produced from known persons or profiles from other crime scene evidence in order to make associations. The high specificity of these associations comes from the frequency statistics associated with each number (or allele) in the profile.[3]

In addition to the nucleus, there is another part of the cell that contains DNA. Mitochondrial DNA (mtDNA) comes from the cell organelles called mitochondria. Mitochondria, and their DNA, are present in human ova and therefore are directly inherited from the mother. The biological father provides no genetic information to the mtDNA. Each mitochondrion[4] contains a copy of its own DNA (mtDNA). One cell may have hundreds of mitochondria, the kidney-shaped units that supply a cell with energy, and therefore have hundreds of copies of the mitochondrial DNA molecule. Unlike the DNA from the nucleus, mtDNA is circular and used by the cell to produce tools needed for the mitochondria to function. Because a person inherits mitochondria directly from their mother's egg, each person born of the same mother has the same mtDNA.

Mitochondrial DNA is important forensically because it can be found, intact, in biological materials where nuclear DNA has degraded. This makes it a valuable tool in any investigation where heavily degraded remains or limited biological materials (i.e., hairs with no root) are collected. Forensic analysis of mtDNA results in the detection of the actual base-pair sequence (A, T, C, G) for two or more regions of the mtDNA. The regions of the mtDNA genome that contain these non-coding, variable regions are called hypervariable (HV) regions. The process of collecting, extracting and producing the mtDNA sequence information from these regions (HV1, HV2, HV3) is time and labor intensive. In many criminal or missing persons cases involving old and degraded human remains, this analysis is the only means of obtaining information to assist in establishing identity.[5]

1.1.2 How DNA Evidence Is Used in Criminal Investigations

DNA analysis in criminal investigations requires the collection of biological evidence at crime scenes and from known persons followed by submission of the evidence to local, state, or federal crime laboratories for analysis by trained forensic scientists. Following analysis, profiles of unknown origin may then be compared to profiles from known persons developed

[3] For additional information about DNA profiles and STRs, please visit http://www.dna.gov/basics/analysis/str.

[4] The singular form of mitochondria.

[5] For additional information about mitochondrial DNA, please visit http://www.dna.gov/basics/analysis/mitochondrial.

in the laboratory or submitted to a database within the Combined DNA Index System (CODIS), the national DNA database system, where they are compared to other profiles from known offenders, arrestees, and other crime scene items. This comparison may occur at the local, state, and federal level. If a profile from an unknown source matches a known individual, notification of the association—more commonly referred to as a "hit"—will be provided to the originating crime laboratory, which will in turn confirm the match and then report findings to local police investigators and/or prosecutors for follow-up investigative purposes. Any association made via DNA, whether through a direct comparison at the lab or the CODIS database, is an investigative aid, not proof of guilt or innocence. As DNA databases expand and DNA collection and analysis techniques improve, the utility of such evidence to the criminal investigator and prosecutor grows. This growth in utility has been accompanied by an ever-increasing growth in demand.

DNA evidence can be classified into two main categories: DNA where the source is known and DNA where the source is unknown. The former is collected directly from persons of interest (e.g., suspects, victims, and/or consensual partners), and the latter may be collected from any myriad of materials and surfaces associated with a crime. Both types of samples are analyzed using the same general processing steps shown in Figure 1. However, because DNA from known persons is collected in a controlled manner (usually a swab or blood sample) these samples are easier to handle and usually produce easily analyzed single-source profiles. Processing DNA from unknown contributors is more time intensive as the analyst may need to search numerous and/or large items (e.g., a bed sheet) for biological stains. These samples are more likely to contain DNA mixtures, which require more time to analyze and interpret. As a result, the processing of DNA from known persons lends itself more readily to automation.

Forensic processing of DNA utilizes predefined, specific locations on the human genome that are non-coding and therefore do not influence a person's physical or biological traits. Therefore, data produced through forensic DNA processing does not reveal any expressed genetic information or physical characteristics of a person. It merely acts as an identifying mechanism that forensic scientists use to determine if there is an association between the evidence sample and a particular person. Once these associations are made and quantified, they are used by law enforcement agencies to aid investigations. The impact of the DNA association on any criminal investigation is dependent on the probative value of the evidence and the context of the investigation.

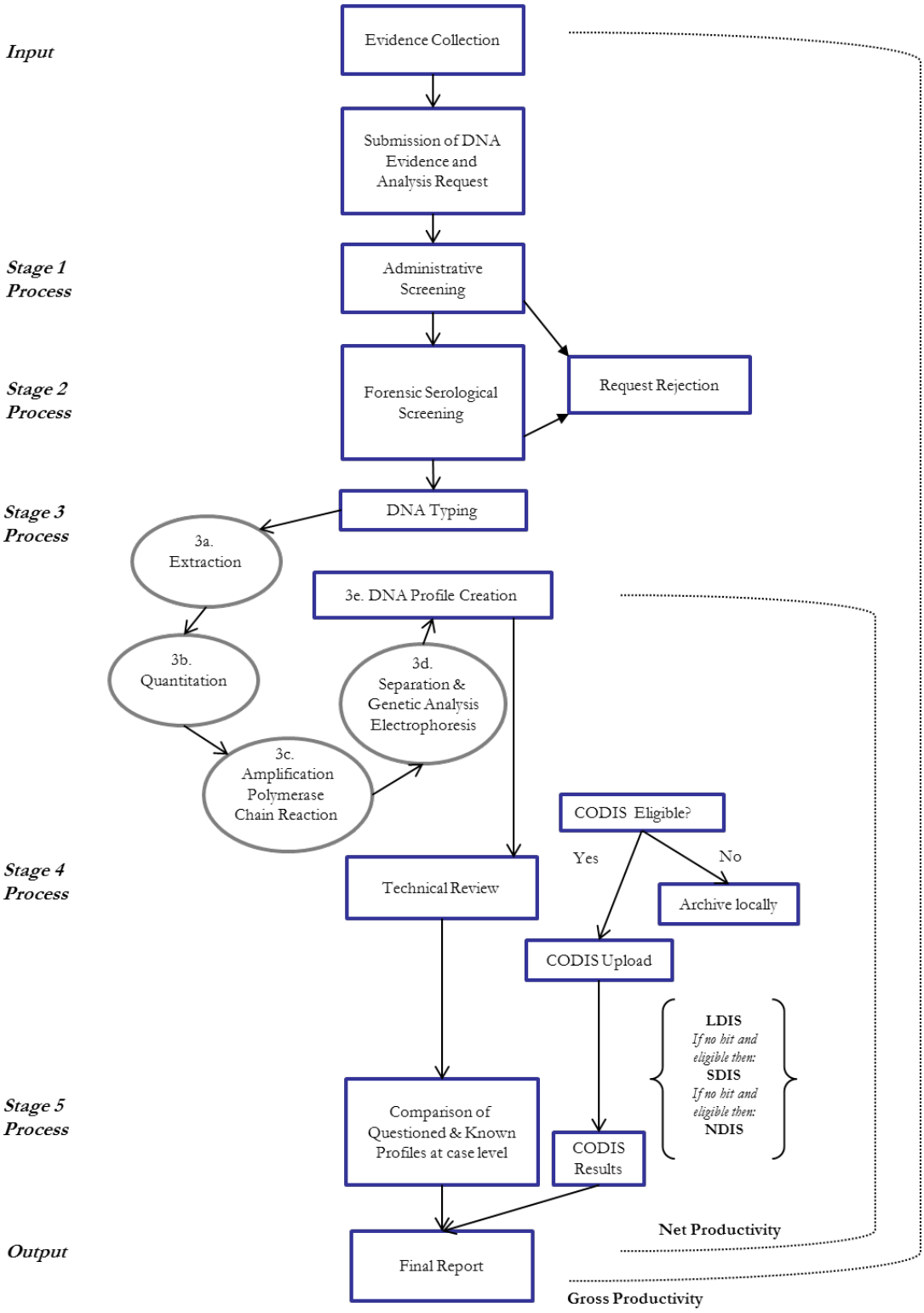
1.2 Processing DNA Evidence by Forensic Crime Laboratories

Figure 1 illustrates the general stages of DNA evidence collection and processing. Since many of the interventions proposed and adopted by the crime laboratory efficiency sites address one or more of these stages, it is important to present a review for reader appreciation

and understanding. After physical evidence is collected and submitted to the lab, it will be assigned to an analyst for processing. In most cases, the first stage in this process is the serological screening of the evidence.

Serological screening is the examination of submitted evidence items for stains or other biological materials. Serologists use chemical screening tests to identify the type of suspected physiological fluids (e.g., blood, semen, saliva, etc.) and to provide investigative information. During serological screening, examiners may recover non-biological trace evidence items (e.g., glass, fibers, powders, etc.) and send them to other units in the laboratory. After the suspected biological materials are identified, cuttings or swabs will be forwarded to the DNA processing unit for analysis, the first step of which is DNA extraction.

Figure 1. Basic Crime Lab DNA Processing Flow



During the extraction step, several types of chemistries and procedures may be used to isolate the DNA.[6] First, the cells are opened to release the DNA-containing material. Second, the proteins that protect the DNA are disrupted to further isolate the materials. Third, the DNA is physically separated from other cellular material and any substance that may interfere with the polymerase chain reaction (PCR) process—that is, the reaction that will make multiple copies of the recovered DNA.

It is important that the quantity of DNA subjected to the PCR is controlled in order to obtain quality DNA profiles. The quantification step determines the amount of DNA in the extracted sample. If that amount is outside the optimal range for PCR, the concentration is normalized by dilution of large amounts or concentration of small amounts of extracted DNA. After extraction, quantification, and normalization, recovered DNA is amplified through the PCR process. This reaction amplifies DNA by making multiple copies of DNA STRs at specific loci in the extracted sample. The materials and chemistries used for this process are usually purchased as complete kits from manufacturers. After amplification, the fragments of DNA are separated via electrophoresis, the electrophoretic data are interpreted, and the DNA profile is determined.

DNA profiles developed from evidence may be entered into a DNA database that is searchable only for law enforcement purposes. These databases exist at the local (LDIS), state (SDIS), and national (NDIS) level, and the software program that coordinates this entire system is called the Combined DNA Index System, or CODIS. While CODIS is the name of the software platform, it has become the de facto name for the databases themselves. Each level contains separate indices for DNA profiles from convicted offenders/ arrestees and forensic samples.[7] As of September 2011, NDIS contained 10,194,686 profiles in the offender index and 365,105 profiles in the forensic index. These profiles have resulted in more than 161,100 hits and 155,100 investigations aided.[8]

1.3 The DNA Processing Backlog Problem

As Roman and colleagues (2008) point out, DNA evidence has become a "widely accepted investigative tool and is routinely collected and analyzed in homicide and sexual assault cases." In its Census of Publicly Funded Forensic Crime Laboratories, the Bureau of Justice Statistics reports that nationwide crime laboratories received over 67,000 DNA processing requests in 2005—over 11 percent more than was reported in 2002 (Durose 2008; Peterson

[6] These chemistries are usually purchased as complete kits from specific manufactures. Several of the study sites proposed to adopt new extraction kits to increase their processing efficiency.

[7] The forensic index contains profiles developed from physical evidence where the source of the profile is unknown (e.g., a DNA profile developed from a blood stain collected at the scene).

[8] http://www.fbi.gov/about-us/lab/codis/ndis-statistics, accessed 19-Oct-2011.

and Hickman 2005). The demand for DNA processing has been affected not only by the nearly ubiquitous use of DNA analysis to aid investigation of serious crimes, but also by expansion of the DNA databases. All 50 states mandate DNA collection from convicted felons and some misdemeanants, and 23 states currently collect from some categories of arrestees. For example, all adult felony arrestees are subject to DNA collection, typing, and profile storage in California (Dale, Greenspan, and Orokos 2006). While the analysis of DNA evidence collected during routine casework and DNA collected for inclusion into the database are usually conducted by separate laboratory sections, the continued growth of the database inclusion criteria requires valuable laboratory resources.

Additional applications of DNA typing, such as identification of unknown deceased persons, missing persons investigations, and the investigation of property crimes, have contributed to the increasing number of DNA analysis requests to forensic laboratories. Unfortunately, this rapid growth in the law enforcement requests for DNA processing has exceeded the ability of crime laboratories to process the growing volume of samples. This has resulted in significant processing delays, increases in turnaround time, and growing backlogs of untested DNA samples, both within laboratories and within police departments that have not submitted samples to labs for processing.

The first systematic assessment of these problems was conducted by NIJ in 2002. At the direction of the U.S. attorney general, NIJ was charged with determining the reasons for the existing DNA evidence processing delays and making recommendations for a national strategy to eliminate unacceptable delays (NIJ 2003). NIJ convened a task force of criminal justice and forensic science experts to achieve its goals. The task force concluded that the processing delays were attributable to the "massive demand for DNA analysis without a corresponding growth in laboratory capacity" (NIJ 2003). It was then estimated that approximately 350,000 rape and homicide cases were awaiting DNA evidence analysis. In addition, it was reported that of the samples awaiting processing, an estimated 90 percent were still in the possession of law enforcement agencies and had not actually been submitted to labs for testing.

The reasons for these backups in processing were identified as primarily being the lack of sufficient evidence storage space, insufficient numbers of trained forensic scientists, and inadequate resources to expand staffing. Moreover, inadequate forensic science curricula and an insufficient number of college programs limited the pool of available forensic scientists. Staff turnover was also identified as a problem. In particular, it was observed that private labs, where forensic DNA analysis is sometimes outsourced, were able to attract experienced examiners with higher salaries. Other resource deficiencies included insufficient lab infrastructure, limited equipment and supplies (and funds to procure them), and a lack of physical space. Recommendations were to 1) improve crime lab capacity, 2) provide financial assistance to build enhanced capacity, 3) eliminate convicted offender DNA

backlogs, 4) support training and education for forensics scientists, 5) provide training for criminal justice and other professionals, and 6) support DNA research and development (NIJ 2003).

These recommendations were incorporated into the President's Initiative Advancing Justice Through DNA Technology, also known as the President's DNA Initiative, which began in 2003 (NCJRS 2008). A variety of new federally funded grant programs were subsequently implemented consistent with the recommendations of the NIJ task force. These included the former DNA Capacity Enhancement Program (2004–2006) and the Forensic DNA Backlog Reduction Program (2005 to present), which were merged several years ago (www.dna.gov 2008). The goal of the DNA Backlog Reduction Program is to reduce the backlog of untested evidence and increase capacity at public DNA forensic laboratories. A primary means is by providing support for capacity increases by funding infrastructure improvements, information management systems, operations, automation, and improved evidence storage (NIJ 2008). As a result of the DNA Backlog Reduction Program, evidence from 135,753 cases has been removed from laboratory backlogs since 2005 (Nelson 2010). While these gains in capacity are impressive, they do not exceed the gains in demands. As a result, evidence backlogs persist and appear to be growing.

It was recently recognized by NIJ that some labs have found other methods to improve production besides simply increasing capacity. These include process mapping, efficiency forums, and business process management models, which are designed to use resources in more efficient ways. To test the utility of these new DNA processing methods, NIJ issued the Forensic DNA Unit Efficiency Improvement Program solicitation in May 2008.

1.4 The NIJ Forensic DNA Unit Efficiency Improvement Program

The NIJ Forensic DNA Unit Efficiency Improvement program originated from NIJ's experiences with the forensic community. NIJ reported that during a grant monitoring meeting, a grantee expressed the importance of developing an integrated approach to multiple bottlenecks in the DNA evidence processing, rather than single-factor advancements, such as more personnel or new equipment. This holistic approach resonated with NIJ program management staff and inspired the creation of this program. This program was specifically designed to *not* be a capacity-building approach to improving evidence processing. Indeed, some of the previous Backlog Reduction program's allowable costs were excluded in this new program's solicitation (e.g., personnel costs to process and analyze casework, purchase of equipment and supplies as stand-alone requests, etc.) (NIJ 2008). Instead, the program provides funding to forensic labs to improve the efficiency of DNA processing through novel approaches.

NIJ issued the solicitation for the Forensic DNA Unit Efficiency Improvement Program in May 2008. Thirteen laboratories applied to the program, although numerous labs

reportedly contacted NIJ about the solicitation after its release. Interviewed NIJ program staff believed the 25 percent match, emphasis on innovative approaches, and the requirement to participate in an external evaluation served as limiting factors for many other interested labs.

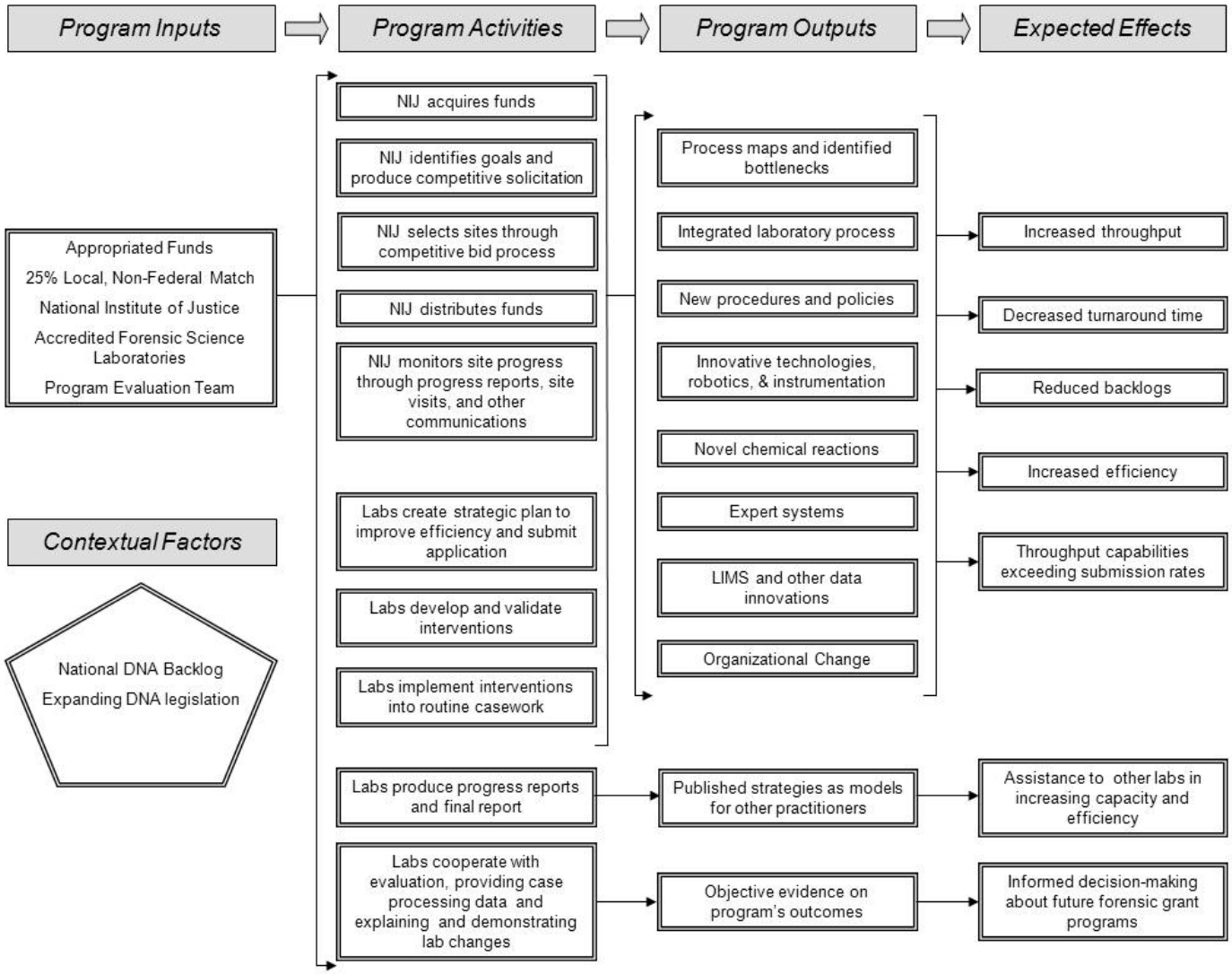
The FY 2008 Forensic DNA Unit Efficiency Improvement demonstration program is administered by the NIJ's Office of Investigative and Forensic Sciences (OIFS). [9] The NIJ is the research, development, and evaluation arm of the U.S. Department of Justice and provides financial support for crime and justice research. The OIFS focuses on research development to support law enforcement and crime laboratories.

Through the Forensic DNA Unit Efficiency Improvement Program, six laboratories were selected to receive funding for innovative and integrated approaches to improve the overall efficiency of DNA evidence processing. In particular, labs were directed to identify bottlenecks in the DNA analysis process and develop cost-effective strategies intended to lead to an improved, more efficient laboratory process. A primary goal of the program was to develop successful, novel efficiency improvement strategies as models for use by other forensic science professionals.

Figure 2 shows a logic model of the demonstration program. Logic models diagram the rationale behind a program, illustrating a series of components that make up the program: (1) *Contextual Factors*, which are important to understanding the external circumstances surrounding a program; (2) *Inputs*, or what resources are needed to begin and continue the program; (3) the *Program Activities* that are performed by various actors of the program; (4) *Program Outputs*, or what is produced by the program; and (5) the *Expected Effects* of the program. This model visually portrays what has been described above about the program and helps guide evaluation and measurement decisions.

[9] At the time of the program's inception, this office was a division within the Office of Science and Technology (OST). During implementation of the Unit Efficiency Improvement Program, the division was moved to an office separate from OST.

Figure 2. NIJ Forensic DNA Unit Efficiency Improvement Program Logic Model



Public forensic science laboratories submitted their proposed strategies to improve efficiency in DNA processing through the competitive solicitation process, and six sites were selected to receive funding. The selected sites, along with brief descriptions of their proposed approach are below:

- Allegheny County Medical Examiner's Office, Forensic Laboratory Division (PA)
    - The site proposed to identify bottlenecks in sexual assault evidence processing through process mapping and then develop a new procedure, which involved using (a) Y-STR analysis and an automated sperm-detection microscope for the purpose of screening items from sexual assault evidence for male DNA, (b) robotics, (c) an expert system for the identification of male profiles and the resolution of mixture samples, and (d) integration of the expert system with the existing laboratory information system.
    - Awarded $382,309 + nonfederal match $127,436 = $509,745 total
- Harris County Medical Examiner's Office Forensic Biology Laboratory (TX)[1]
    - The site proposed to hire Forensic Science Services, LTD, to perform a diagnostic review of the lab workflow and implement recommended changes.
    - Awarded $504,000 + nonfederal match $168,000 = $672,000 total
- Kansas City Police Crime Laboratory (MO)
    - The site proposed to create a more streamlined system for processing items from known sources (i.e., known standards) including implementing (a) a new enzyme extraction method, (b) automated extraction with robotics, (c) standard cutting of items to potentially eliminate quantification steps, (d) automated amplification process with 96 well plates, and (e) an expert system for data interpretation.
    - Awarded $90,000 + nonfederal match $30,000 = $120,000 total
- Louisiana State Police Crime Laboratory (LA)
    - The site proposed to hire a consultant to conduct process mapping and make recommendations. After undergoing the consulting process, the site chose to (a) adopt Lean Six Sigma principles in order to apply

[1] During the program period this facility changed its name to the Harris County Institute of Forensic Sciences. The name presented in this report is the one used in the FY 2008 project proposal from this site.

business management ideas to laboratory tasks, (b) shift non-analysis tasks to non-examiner personnel, (c) outsource robotics validation, (d) document management, and (e) adopt Lean Six Sigma principles to purchasing activities.

  - Awarded $ 450,000 + nonfederal match $150,000 = $600,000 total

- San Francisco Police Department Criminalistics Laboratory (CA)
  - The site proposed to develop a comprehensive case management system with modules for cold hits, mixture interpretation, administrative features, data review, report writing, quality assurance, and grant management.
  - Awarded $1,024,467 + nonfederal match $341,489 = $1,365,956 total
- University of North Texas Center for Human Identification (TX)
  - The site proposed to implement a variety of new approaches to more efficiently analyze mitochondrial DNA family reference samples, including changes related to chemistry, robotics, expert filtering software, and data tracking.
  - Awarded $601,632 + nonfederal match $200,544 = $802,176 total
  - Final period of performance: 10/2008–12/2010

The Harris County Medical Examiner's Office Forensic Biology Laboratory ultimately decided not to participate in the program and declined the grant award during the summer of 2009. The San Francisco Police Department Criminalistics Laboratory withdrew from the program in June 2010.

The original award period for the Forensic DNA Unit Efficiency Improvement Program was October 1, 2008, to March 31, 2010, although sites received multiple no-cost extensions due to implementation delays. As will be noted in the individual case studies, by the end of the evaluation in 2011, components of a number of programs had yet to become operational.

1.5 Productivity, Capacity, and Efficiency

In order to better understand the nature of the problem targeted by the Forensic DNA Unit Efficiency Improvement initiative and the current evaluation methodology, clear distinctions must be made between a number of terms and concepts, which appear to be used interchangeably in much of the extant literature. The NIJ defines backlog as a case "that has not been tested 30 days after it was submitted to the laboratory" (Nelson 2010). This type of comparison is one simple indicator of *productivity*—the difference between output and input over a given time interval. So, for example, in 2005 crime laboratories began the year with a

backlog of 24,030 requests, received 67,009 new ones, processed 52,812 of the total, and ended the year with 38,227 in backlog (BJS 2008). Therefore, backlog productivity fell short by 14,197 requests, which could also be represented proportionally as a 59 percent backlog increase. While productivity may be important, it holds limited evaluative significance in and of itself; it is simply a measure of output to input.[2]

Past literature has suggested that one means by which to change crime lab productivity is to increase *capacity*. This refers to bringing more resources to bear on the production process (and is the method used in the DNA Backlog Reduction Program). The underlying assumption is there is a direct positive relationship between capacity size—such as the physical space, number of trained personnel, instruments and tools (robotics, for example), support staff, and other resources—and productivity. Therefore, if the average forensic examiner processes 77 DNA requests in a year, as they reportedly did in 2005 (BJS 2008), then full productivity can be achieved by adding the requisite number of new examiners and support resources. This is in fact what BJS suggests by calculating a necessary 74 percent increase in examiners required to eliminate the 2005 backlog (BJS 2008).

Efficiency, on the other hand is a comparative term for the measurement of productivity in relation to a particular resource unit (Heyne n.d.). Using the above example, if one could increase the number of samples each current examiner could complete from 77 to 133 a year through new processes or technology, productivity would increase. The change in productivity however, would be the result in this example of efficiency, not capacity. The primary goal of this initiative is increasing efficiency, not just capacity, so each of their relative effects on processing will need to be carefully distinguished for evaluation purposes.

The evaluation of this initiative becomes more complex when one considers the step-by-step sequential processing requirements (both legal and procedural according to accrediting standards) associated with DNA evidence analysis. A basic generic representation of the DNA lab processing flow stages is presented in figure 1. As can be seen in this simple representation, there are several intermediary processing stages between evidence submission (input) and reporting of results (output). Of particular note is stage 1, administrative screening, when requests can be rejected for a variety of reasons—policy, legal requirements, nature of the evidence, or contextual information about the investigative questions. The volume of rejections is important for measuring productivity. As noted above, changes in yearly backlogs are comparisons between inputs and outputs and therefore *gross productivity* measures. *Net productivity*, on the other hand, is defined as a comparison of input minus request rejections with outputs. Net productivity may also be affected by rejections at stages 2 and 3, where sample quality or quantity may be deemed as less than robust enough for

[2] As such, productivity can be increased simply by restricting or reducing inputs, not just improving processing volume.

typing. Further, items may be withdrawn from processing from stage to stage for a variety of reasons—for example, the closing of a case or the generation of a poor-quality DNA profile not suitable for comparison. To illustrate this point, in the recent UI study of the use of DNA in property crimes, it was found that evidence from 70.3 percent of cases yielded a profile, while only 54.7 percent of cases had a profile uploaded to CODIS (Roman et al. 2008).

2. EVALUATION METHODOLOGY

2.1 Goals and Objectives

The Urban Institute's primary goal was to systematically evaluate the implementation and potential outcomes of what NIJ considered to be novel and innovative approaches to improving DNA crime laboratory efficiency. This was accomplished through two primary research methodologies. The first was documenting the implementation of the Forensic DNA Unit Efficiency grants at each of the funded sites through a case study approach. The second was to assess possible productivity and efficiency outcomes at the study sites within the context of implementation milestones. Both longitudinal and pre/post research designs were utilized for this objective.

It is important to note that this evaluation does not examine the effectiveness of individual interventions, nor their performance or validation testing in controlled experimental laboratory settings. Please see the individual sites' published final reports to NIJ for internal validation findings and other related information (listed in appendix B). Instead, the current study asks the question: What were the observable effects of the NIJ Forensic DNA Unit Efficiency Improvement Program in the real-life settings and functions of laboratories? This is an important distinction, because it is quite possible that a developed intervention could show time savings in a controlled experiment but then not impact overall turnaround time (e.g., if those time savings are then lost in other delays between stages or while waiting for one's turn on a laboratory instrument).

The researchers at the Urban Institute gathered information between January 2009 and September 2011 from multiple sources to achieve its evaluation goal and objectives. Some of these data included NIJ program materials, interviews with program administrators and laboratory staff, on-site observations, and data from laboratory information management systems. More details concerning the kinds and quality of data used for both the process and outcome components are discussed in more detail in methodology descriptions and elaborated on in more detail in the appropriate site case studies and appendices.

The implementation evaluation findings are reported descriptively. Lab productivity and efficiency outcomes were examined using descriptive statistics, longitudinal trend representations, independent sample t-tests, and regression analyses. The UI researchers also tracked other changes and events that took place in individual labs over the course of the study in order to control for potential outcome confounds to the degree possible.

2.2 Process Evaluation Methodology

2.2.1 Implementation Data

In order to document site implementation, including successes and challenges, the research team collected information about the NIJ Forensic DNA Unit Efficiency Improvement Program from numerous sources. The first included reviews of written grant development documents, award proposals, process map diagrams, site and consultant progress reports, dissemination materials (such as journal articles and slide presentations), sites' final NIJ technical reports, and similar written materials. UI researchers also conducted interviews with NIJ program administrators about their views of program development, initiative goals, and award decisions.

Researchers from the UI team also made multiple visits to each site to tour the laboratory, conducted interviews with staff essential to the site's projects, and received demonstrations of the new lab processes, technologies, and other novel approaches that were thought would have an effect on processing efficiencies. Monthly phone calls were also conducted with the lab's primary point of contact to track progress on the grant activities and any other changes or events in the lab that had the potential to impact processing and evaluation findings. In addition, members of the UI team attended national forensic science conferences and symposiums to view site presentations and conduct in-person meetings with key site project staff and program managers from NIJ.

2.2.2. Data Analyses

Implementation data from program materials, site project materials, site visit demonstrations, and interviews with NIJ members and participating laboratory staff were synthesized to understand the project goals and outcomes and the implementation process and challenges, both for the program as a whole and for individual sites. This information was then used to produce logic models, grant milestone timelines, and descriptions of each site's implementation for this final report. The draft site logic models were reviewed by NIJ and each site upon completion and revised as necessary based upon their feedback. The logic models were used to facilitate understanding of each site's grant goals and to help identify outcome measures for each site and guide the evaluation.

2.3 Outcome Evaluation Methodology

2.3.1 Outcome Data

In order to assess changes in DNA processing, the UI team collected detailed case processing data from each laboratory. Case and/or sample processing data were extracted from each

lab's LIMS[3] for all DNA cases or requests received between January 1, 2007, and January 31, 2011. While the research team requested the entirety of the lab's database for analysis (with the exception of identifying information), some labs were only able to extract specific variable data. Available data fields and quality of data tracking varied substantially across labs, and labs differed on what unit of analysis (i.e., case, sample) they used to track casework. At minimum the following fields were collected across all sites:

- Case or sample ID
- Date of case/sample/routing start point and end point
- Dates for intermediary stages
- Item or sample descriptions
- Criminal offense
- Analyst responsible for case/sample/routing
- Priority designation or other indicators of potential prioritization (e.g., suspect present)

Resource indicators were also collected to assess stage and overall unit *efficiencies*, the ultimate goal of this evaluation. Measures of resources served as the denominators of the efficiency ratio indices (productivities were the numerators).[4] For example, the number of samples processed each month (productivity numerator) divided by the annual budget expenditures (unit resource unit denominator) could be used as one efficiency ratio index. Each site provided the number of analysts and technical support staff employed in the DNA (and serology for Allegheny County) unit by year. Part-time status and partial-year employment were accounted for if the data were detailed enough to permit such calculations.[5] All but one site were able to provide operating budget expenditure information[6] for the DNA unit, but the form of this data varied across sites (see table 1). Because monthly-level data were not available, the research team had to assume that expenditures and labor

[3] We refer to any laboratory's electronic data management system as a LIMS, although it may not be a formal laboratory information management system, such as fully integrated commercial software designed to track processing and produce reports. For instance, one lab tracked information in Excel spreadsheets.

[4] Please refer to the previous definitional distinctions made between productivity and efficiency for an understanding of the importance of resource indicators as denominators for measuring efficiency.

[5] If staff members were present for only part of the year, they were counted as the proportion of months they were employed (i.e., if someone left in May 2010, he or she would be counted as 5/12 of a FTE employee for that year). Part-time employees were similarly counted as partial staff counts, calculated based on the number of hours worked per week.

[6] Sites were unable to provide reliable grant expenditure information by month or by year.

were fairly constant across the year. Each site except UNT, which did not provide budgetary information, had both a financial efficiency index and a labor efficiency index for each outcome: throughput and turnaround time. These efficiency indices were calculated as described earlier with the productivity measure (throughput or turnaround time) divided by the resource indicator (annual labor counts or budget expenditures). Greater efficiency is evident from *higher* throughput efficiency indices and *lower* turnaround time efficiency indices.

Table 1. Data Characteristics by Site

Site	Unit of Analysis	Sample	TAT Definitions	Intermediary Stages	Resource Indicators	Data Issues
PA	Serology case (includes DNA work, if applicable)	All Serology/ DNA forensic evidence *(N = 1,518)* *(N < 400 for cases with DNA)* Subsample: sexual vs. non-sexual offense	Serology: Assign to administrative review DNA: Extract to administrative review	1. Serology submission 2. Serology assignment 3. Serology report 4. Serology administrative review 5. DNA submission 6. DNA assignment 7. Extraction 8. Quantification 9. Amplification 10. CE injection 11. DNA report 12. DNA administrative review	• Annual number of serology and DNA analysts • Annual operating budget expenditures (primarily supplies costs)	• Small number (< 400) with DNA work done or requested • No indicator for canceled status • Reported inconsistent data recording practices
KS City	DNA sample	All DNA forensic samples *(N = 10,296)* Subsample: known standard vs. unknown	Assignment to report	1. Submission 2. Assignment 3. Extraction 4. Quantification 5. Amplification 6. CE injection 7. Analysis 8. Technical review 9. Report	• Annual number of DNA analysts • Annual (fiscal year) supply and equipment expenditures	• Changed LIMS • Stopped casework in March while transitioning LIMS • Chronological order issues • No indicator for canceled cases

Site	Unit of Analysis	Sample	TAT Definitions	Intermediary Stages	Resource Indicators	Data Issues
UNT	Routing, Case	All family reference samples (FRS) *(Rout = 3,429 Case = 3,079)* Subsample: mitochondrial vs. non-mitochondrial FRS	Date "started" to technical review	1. Date "started" (when analyst first handles evidence) 2. Extraction 3. Analysis 4. Technical review 5. Report 6. Administrative review 7. CODIS entry 8. CODIS notification 9. Date "testing completed"	• Annual number of analysts qualified for family reference sample work	• Changed LIMS during study period • Report not completed unless match made to missing person • Separate timed experiment
LA	Case	All DNA forensic evidence *(N = 4,325)* Subsample: outsourced vs. non-outsourced	Assignment to administrative review	1. Offense date 2. Request 3. Assignment 4. Report 5. Technical review 6. Administrative review	• Annual number of trained, caseworking analysts • Annual (fiscal year) state budget expenditures for DNA Forensic Unit	• Few intermediary stages • Few predictor variables

Note: No budget information includes grant expenditures because sites could not easily produce this information by year.

Measurement Definitions and Complications

Before describing the analyses used, a few notes must be made about the definitions used for the study's performance measures and the general nature of DNA processing and turnaround times. One challenging aspect of the current study's methodology was to produce suitable outcome definitions. Based on the UI team's prior forensic experience, consultations with the team's internal and external forensic experts, and interactions with each of the sites, the following performance measures were chosen:

- Throughput = # of samples/cases/routings completed[16] per month
- Turnaround time (TAT) = # of days between start and end points for entire (or stage of) DNA processing/analysis
- Throughput/staff = throughput divided by number of relevant staff
- Throughput/budget = throughput divided by relevant budget expenditure amount in dollars
- TAT/labor = turnaround time divided by number of relevant staff
- TAT/budget = turnaround time divided by relevant budget expenditure amount in dollars

While the research team strove for comparable measures across the sites, this ultimately did not prove feasible. This was primarily due to the fact that the availability, content, and quality of these data varied significantly across individual laboratories. Each site had different definitions of overall turnaround time, available stage-level turnaround times, and resource indicator measures (i.e., staff and budget) (see table 1). Because of these differences, cross-site comparisons were limited. Nonetheless, the definitions were consistent across time *within* each site to allow for valid comparisons before and after the grant activities.

In defining turnaround time, the research team attempted to use the most complete measurement of turnaround time possible at each site. However, it should be pointed out that submission dates were not used as the start point for defining turnaround time. While the submission of a request might appear to be a sensible beginning point for defining turnaround time, long wait periods between submission and assignment could mask any efficiency gains obtained at later points in the process. The time between submission and assignment can also be affected by many factors not relevant to the interventions, such as waiting for additional

[16] Canceled cases are not included. A case is considered completed if it had a "complete" status (not available in all sites) or had a date listed for the technical review, report, or administrative review (dates used depend on what stages site tracked; see table 1).

evidence or trial dates. Therefore, when available the date of assignment was instead used as the preferred start point for processing and analysis.[17] The end date was the date of report completion or review, depending on which of these stages occurred later and had more complete data.[18] Because none of the sites consistently recorded information about *time* of completion in their LIMS, turnaround time could only be calculated in number of days. This more gross measure of turnaround time could also mask efficiency gains of smaller magnitude than a single day.

A second important analytic issue involves the nature of DNA processing and the chronological ordering of stages. A sample may require a rerun at various stages. For instance, a sample may need to be re-extracted if analysts cannot find DNA after the quantification stage. A sample may be reinjected (or even return to the extraction stage) if the data derived from this first capillary electrophoresis run is not usable. Depending on how dates are recorded in a laboratory's LIMS, this may produce an irregular chronological progression of dates (e.g., analysts overwrite the extraction date with the re-extraction date but leave all other dates the same, giving an extraction date that falls later than the quantification, amplification, and injection stage dates).

For case-based tracking, a case is typically not considered complete until all relevant samples are finished. Apart from the fact that multiple samples are often submitted at once (and all of these would need to be completed before a report was written), a case may also involve multiple submissions across time (e.g., a suspect's reference sample is submitted to the lab months later) or forensic work performed in other departments (e.g., a firearms request in addition to a DNA request from the same crime scene). The same samples may even be resubmitted at a later date if a court requires a reanalysis or additional analytic work in preparation for a trial.

Due to these two events (reruns of samples or multiple/resubmissions of samples for a case), the sequence of dates as a sample/case progresses through analysis is not as linear as

[17] Allegheny County's DNA submission and assignment dates were not usable because the site reported that analysts frequently did not enter this data accurately. Instead, the DNA extraction date was used as the start point for DNA analysis (serology assignment date was used as the start point for the serology component of the case). Because Allegheny County did not plan any interventions for the extraction stage (and therefore we would not expect the grant to impact the amount of time between assignment and extraction completion), this was deemed a suitable solution. UNT did not record assignment date in its LIMS. Instead, the date was used for "date started," a somewhat nebulously defined stage in the site's LIMS which more or less correlates to when the samples are first handled by the analyst.

[18] Administrative review was used as the final end stage for Allegheny County and Louisiana. Kansas City did not report dates for the administrative review, so the date of the report was used (this was later than the technical review date). UNT had very few report dates within the database because the analysis of family reference samples does not result in a report unless and until there is a match with a missing person or unidentified human remains. Instead, the technical review date was used.

might be expected. This creates difficulty in the estimation of turnaround times and assessing the impacts of the grant program. Any efficiency gains might be masked by lengthy wait times between submissions, and stage-level turnaround times may seem exceedingly large for stages where reruns were performed. The discussion below describes how these issues were addressed in each site depending on the structure of each site's data.

Structure of Analytic Files

Each site had a different unit of analysis structure (see table 1) and amounts of information available about cases. The structure of each site's data had significant implications for the interpretation and ability of the present study's analyses to detect changes across time. In addition, analyses were run for different selections of data, depending on the site (e.g., only known standards, only family reference samples). The analytic series used are shown in table 2.

Allegheny County tracked its DNA workload information by the "serology case" unit of analysis. All serology cases were included, because proposed grant activities targeted both the serology and DNA analysis stages. The lab reported that they considered every serology case to have the potential for DNA analysis. However, only a small portion (24 percent) of serology cases across the period of the evaluation progressed to DNA work. Because Allegheny County tracks by *case*, analyses were necessarily limited in their ability to detect changes in turnaround time (see section 2.5, Evaluation Challenges and Limitations). The research team originally intended to focus on sexual assault offenses for Allegheny County. However, these cases comprised too small a percentage of both serology cases (22.5%, $N = 252$) and as well as those involving DNA (27.7%, $N = 107$). Instead, UI included all forensic serology and DNA casework, which aligns with the site's assertion that all serology samples are considered contenders for future DNA work.

In addition, this site reported inconsistent tracking of information about multiple submissions and rerunning of samples within its LIMS during the study period (staff have since made performance measurement a high priority and are now reportedly using better data-recording practices). The site did not always track multiple submissions or supplemental work in the databases made available to the UI research team,[19] and when staff did record this information, it was done inconsistently (i.e., it might have been given a new entry with the same case number, they might have overwritten the dates within the existing case entry, or they might have documented that supplemental exams or multiple extractions were completed). Because of these inconsistencies, the team could not always determine with confidence whether a case included multiple submissions or if a sample was rerun at one of the stages. If there was an indication that additional work was done, it was not always clear if

[19] This information was typically stored in other locations, such as hard copy case files. However, the UI researchers were unable to feasibly access this information.

this work was due to a new submission or due to reworking of the existing samples, or when this work occurred. Given the state of these data, the research team had to account for these situations by creating a variable that identified cases that had *some* evidence of additional work done (either through multiple submissions or through reruns). Analyses of case turnaround times control for this additional work and the expected longer turnaround times for these types of cases. However, because some situations were not tracked or dates may have been merely overwritten with the most recent dates for additional work completed, it is likely that there are cases in the database for which multiple rounds of work were completed, but the case could not be identified as such.

Kansas City's data consisted of large numbers of separate databases with different units of analysis (item vs. sample vs. case) that required merging. Kansas City was the only site that tracked by DNA sample. However, dates for some stages were only available at the case-level. These dates have the same interpretation challenges discussed above, where efficiency gains could be masked at the stage-level. In particular, the technical review and report stages were only available at the case-level in their new LIMS, so overall turnaround time and stage-level turnaround times from the data analysis/interpretation stage forward were not very sensitive to grant outcomes after April, 2010. Case-level interpretation, technical review, and report dates were also used when sample-level analysis dates were missing in their Excel data. Because Kansas City was more consistent in their electronic tracking practices, the research team was able to create separate variables to flag when a sample experienced a rerun[20] or if a case-based stage date was associated with a case that had multiple submissions (and therefore would be expected to have a longer turnaround time). The research team also has more confidence in the validity of these flags, as Kansas City was more consistent in its methods of tracking this information. Kansas City's grant focused on the processing of known standards, so analyses narrowed in on this type of casework in order to enhance the ability to detect efficiency gains which might be masked by analyzing the entire dataset.

The University of North Texas also changed database systems during the data evaluation period. However, all of its data were transferred into its new LIMS, allowing for easy comparison across time. The new LIMS produced multiple tables that tracked information separately by "routing" and by case. A routing was defined by the site as any time a sample or set of samples underwent the DNA process. This construct is somewhat comparable to the idea of a submission or request. While UNT's grant focused particularly on the processing of mitochondrial family reference samples, the single implemented intervention was expected to improve efficiency for both mitochondrial and non-mitochondrial family reference samples.

[20] Reruns at all stages are tracked in the new LIMS, but only reinjections during the capillary electrophoresis stage were tracked in the Excel data.

Therefore, the research team explored the grant's effects on all family reference routings, regardless of whether they involved mitochondrial DNA.

A routing could include single or multiple samples, but would always be for a single case (unlike tracking by a batch). New submissions for the same case would result in a new routing; however, reruns would remain in the same routing. Reruns were not identified in any way in the LIMS data; if a sample was rerun at a certain stage, the first date would be overwritten by the date when all reruns were complete. Therefore, analyses cannot account for extra time spent in reruns for UNT. However, the UNT data do allow for a more precise understanding of whether a case had multiple submissions through its division of data by both case and routing. Primary analyses use routing as the unit of analysis, as this measure would be most likely to accurately detect changes in turnaround time due to grant activities.

The final site, Louisiana, had a single database that did not require merging. Louisiana's data were quite straightforward, although they lacked much of the information available at other sites. There were no intermediary stages reported between assignment and completion of a draft report. There were also limited variables tracking other case characteristics that might help to predict a case's turnaround time. Louisiana tracked information by the case unit of analysis, and the grant targeted all forensic (i.e., not convicted offender/arrestee samples) casework. LIMS data offered no indication of whether cases had multiple submissions or reruns. Consequently, this could not be controlled for in analyses.

Table 2. Outcome Series and Analytic Variables

Site	Analytic Series	Pre/Post Intervention Point	Secondary Intervention Dates for Regression	Regression Control Variables
Allegheny	Serology cases (*N* = 1,511) DNA cases (*N* = 365)	N/A	11/2010: DNA LIMS Module (*STaCS*)	Other events: lab move, summer internships, new quantification instrument use Rerun or multiple submissions # Serology items Item type Known suspect Offense Analyst experience
Kansas City	Known standard	3/2009: New extraction	10/2009: Robotics &	Other events: Identifiler Kit,*

Site	Analytic Series	Pre/Post Intervention Point	Secondary Intervention Dates for Regression	Regression Control Variables
	samples (*N* = 3,173)	method for buccal swabs	new extraction method for blood samples 12/2009: Standard cutting 7/2010: Robotics	Updated CE instrument,* LIMS transition Rerun or multiple submissions Type of standard Analyst experience
UNT	Family reference routings (*N* = 3,428)	3/2009: Auto-fill worksheets	N/A	Priority status Mitochondrial analysis Analyst experience
Louisiana	Non-outsourced forensic DNA cases (*N* = 2,748)	8/2010: LSS process piloted	N/A	Other events: outsourcing, electronic logs, tandem teams, lab expansion, outsourced training, new computer/scan equipment, business unit Number of items Item type Offense Analyst experience

* Some "other events" occurred simultaneously with intervention implementation and are therefore captured in the same variable as the simultaneously implemented intervention.

2.3.2 Data Analyses

Case processing data were used to measure the impacts of the program on the productivity and efficiency of DNA processing within each lab. In order to assess these outcomes, the UI research team conducted four sets of analyses for each site: (1) descriptive, (2) pre/post comparison tests, (3) longitudinal trend representations, and (4) regression analyses. The methodology for each of these is described below. Raw data were transformed in two primary ways to facilitate analyses, including (1) conducting median imputation of turnaround times for unfinished cases to prevent skewed results, and (2) addressing extreme

outliers and other data anomalies. The justification for and methodology of these transformations are detailed in appendix B.

Descriptive Analyses

Descriptive outcome statistics were examined and reported for the overall sample (i.e., across all months) at each site, including mean, median, standard deviation, range, and skew. These statistics were calculated for both the original, raw data and the working data, which underwent median imputation and cleaning for data anomalies (see appendix B). Specifically, these statistics were reported for each site's (1) throughput, (2) overall case turnaround time, (3) stage-level turnaround times, and (4) each of the previous listed measures divided separately by both the staff and budgetary resource indicators (resulting in the efficiency ratio indices).

Comparison Pre/Post Analyses

Each site with at least one implemented intervention underwent analyses to determine whether the average number of cases completed per month changed after implementation. In most cases, two similar statistical tests were utilized; independent sample t-tests and the Mann-Whitney U-test. T-tests measure whether differences in means (the sum of observations/number of observations) of outcome measures are statistically significant. On the other hand, the Mann-Whitney U-tests measure whether differences in medians (the middle observation of a rank ordered distribution) of outcome measures are statistically significant.[21] The delineation of pre/post periods was contingent upon each site's implementation timeline. The intervention points used for each site are shown in table 2.

Allegheny County did not experience any stable implementation by the current study's end (see section 3.3, Implementation Findings, for more detail); therefore, no pre/post comparison analyses were completed. Kansas City implemented its grant developments incrementally, making it difficult to pinpoint a single intervention date. The research team chose to use the first implementation milestone, when the new extraction method started being used routinely for known standards casework. This date represents the first major change to Kansas City's known standards processing and marks the beginning point of other future grant activities. University of North Texas was only able to transfer one new development into casework by the time of this report; this change occurred in March 2009 and was expected to influence both mitochondrial and non-mitochondrial family reference sample processing times. Louisiana's method of implementation had analytic advantages, as it developed and planned an entirely new process for analyzing DNA samples over the period of multiple months but implemented the process as a whole starting on July 26, 2010.

[21] Readers unfamiliar with these inferential statistics are referred to P. E. McKnight and J. Najab, "Mann-Whitney U Test," *Corsini Encyclopedia of Psychology* (2010, 1) for more information.

Additional changes to the lab occurred in 2011; however, the bulk of the grant's changes were completed by August 2010, and additional changes after that point mostly occurred after the period for which the team had data (i.e., after February 2011).

Regression Analyses

In order to understand the influences of the grant activities on turnaround time measures, the UI research team used negative binomial regression analyses[22] to predict turnaround time based on intervention implementation, as well as other case/sample/routing characteristics expected to impact the duration of casework. A series of multiple regression analyses was run for the overall turnaround time, as well as stage-level turnaround times, within each site. Multiple regression is an analysis technique which measures the relationship of multiple independent variables to a dependent variable of interest. The estimated effect (regression coefficient "b") of each independent variable is the relationship between that variable and the outcome variable while holding all other independent variables constant. This allows for an estimation of the unique influence of each independent variable.

Negative binomial models were selected for the regression analysis because turnaround time is a "count" of days and the distributions of turnaround time had substantial right/positive skew. The regression coefficient for a negative binomial regression is interpreted as the percentage change in the dependent variable for every one unit increase in the independent variable.[23]

Within each site, regression analyses were conducted with dichotomous (i.e., "dummy") intervention variables that indicated whether the site's intervention had been implemented. In addition to using the intervention milestone dates used in the comparison t-test analyses, secondary "dummy" variables were also created for incremental interventions (see table 2) at Kansas City. Other events not related to the program grants (i.e., new technology from other grants, facility changes, policy changes) expected to impact turnaround time were also included to control for these influences separately from the grant interventions.

Other characteristics of the case/sample/routing were included in analyses as control variables to account for other factors affecting turnaround time. The included control variables depended on available data in each site. The research team reviewed all existing variables in each site's submitted dataset and selected variables that had valid data and were expected to impact turnaround time (see table 2). Unfortunately, many characteristics of interest were not available in the data. Louisiana, in particular, had little information about

[22] Three models for Kansas City used the zero-inflated negative binomial distribution due to the presence of a large number of "0" values for stage-level turnaround times (i.e., stages that occurred within the same day).

[23] Again, readers unfamiliar with this statistical test are referred to J. M. Hilbe, *Negative Binomial Regression* (Cambridge University Press, 2011) for additional information.

case characteristics available in the extracted LIMS data. Other sites had variables of interest, but they had to be excluded due to a large number of missing values or their absence after a change in LIMS. Variations across analysts were examined by including a measure of analyst experience in the regression analyses.[24] Ordinary least squares regression analyses were run before the negative binomial regression analyses to obtain diagnostics on multicollinearity for each model's included variables.[25]

Because Allegheny County did not fully implement any grant-related activities, this site's regression analyses help to elucidate what case characteristics influence turnaround time rather than focusing on assessing the impacts of the grant itself.

2.4 UNT Timed Experiment Supplemental Analysis

UNT was the only site that operated a separate research and development group and a casework unit. All of the new chemistries and technologies were developed, validated, and tested in the research division. During the grant period, only one of the grant interventions, the auto-fill Excel worksheets, was implemented by the casework division. Therefore, it was apparent that the primary data source (LIMS), data type (case processing data), and analytical methodology were not wholly suitable to properly evaluate this site.[26] The UNT staff worked with the research team and NIJ program management to design a timed experiment that collected DNA processing stage time data from both the new method designed by the research division and the method used by the casework division. At the time these data were collected, the new Excel worksheets were already implemented; therefore, any change in work time caused by this intervention was not captured by this exercise.

The research division and the casework division sections were asked to record the date, start time, and end time for each DNA processing substage for several batches of family reference samples. The casework division used the current laboratory procedure, while the research division used the procedure developed through grant activities. Both procedures included STR processing and mtDNA processing. Average DNA processing substage times were calculated from each observation from the time data supplied by the sites. If a particular

[24] Years of experience was determined by calculating the number of years between the time the case/sample/routing was started or assigned and when the employee was hired. An attempt was made to also account for years of experience before hire at the current agency. After consulting the team's expert forensic consultant (who has previously managed a crime laboratory) and inspecting the hiring criteria for open job position announcements, the team decided to add 4.5 years to analysts hired at intermediate levels and 7.5 years to those hired at senior levels.

[25] Multicollinearity is when predictor variables are strongly correlated, which can affect the calculations and validity of individual predictor variables.

[26] However, the analytical methods are appropriate to establish baseline case processing data for the casework division and evaluate the impact of the single implemented intervention, autofillable Excel worksheets.

substage was performed multiple times within a single batch it was treated as a separate observation of the performance of that substage.

The UI research team calculated difference in the averages, which represents time gained or lost with the new research division procedure, for each substage of DNA processing. Results of the timed experiment analysis are presented in section 7.4.4. While the UI research team performed analyses on UNT's data, it should be noted that UI was not involved in the performance of the experiment. Further, results from this experiment only show the *potential* for time-savings changes from UNT's new workflow—it does *not* provide findings on the actual impact on the lab's processing of DNA samples.

2.5 Evaluation Challenges and Limitations

The research team believes the current study is the strongest feasible design given the project goals and constraints of the project timeline and available data at sites. However, there are substantial limitations with the current study's design. These are detailed below.

2.5.1 Evaluation Challenges

The first challenge was the delays encountered by all of the program sites in the implementation of grant funded activities (described in further detail in each site's case study). While the grantees' no-cost extensions were accompanied by extensions for the evaluation, the research team was not permitted to extend the project for as long as they would have preferred. These timing constraints required the research team to collect final data through January 31, 2011, although the grant periods of performance did not end until March 2011. Although the research team continued to document grant activities occurring through March 2011, they could not examine data that might show additional impacts on productivity and efficiency up to this period—or beyond. An intervention's effects are sometimes felt more strongly as time continues, and this was not captured by the current evaluation. Moreover, the short follow-up periods meant there was little post-intervention data on which to draw conclusions about the effects of the grant program.

Another complication of the study design is that there were numerous other events occurring at the site laboratories. These other changes at the lab, including policy changes, personnel turnover, facility expansions, and technology and activities funded by other grants, had the potential to confound the study's results. Nonetheless, the research team did seek to document as many other external events and changes expected to influence DNA processing as possible and controlled for these changes in the regression analyses. Unfortunately, pre/post comparison tests (t-tests and Mann-Whitney U tests) are not able to control for these other events.

An additional challenge of this evaluation pertains to the nature of forensic work. As described above in the Measurement Definitions and Complications section, forensic cases do not always follow a linear path from start to finish. Moreover, DNA casework is highly variable and dependent on the type of case and quality and type of evidence. This makes it difficult to measure changes across time when DNA cases are not identical "cogs" expected to take similar amounts of time to process. One long and complex case could cause a spike in the monthly average turnaround time. An attempt was made to adjust for such situations through the exclusion of large outliers under the assumption that remaining variability across cases was evenly distributed across the time period; however, it was possible that such irregularities could cause bias in the study's results in ways not able to be determined.

The intersection of the forensic and social science worlds is bringing informative new research to both fields. However, there are inherent difficulties in this type of work. For instance, the two fields have different goals and disparate conceptualizations of "research." Individuals from a physical science background are used to thinking about research in very controlled settings. However, social scientists rarely have the fortune of such a study environment. The research team needed to engage in educational efforts to explain the goals, methodology, and limitations of the current project to partners and stakeholders (see section 2.1, Goals and Objectives, for more detail on the current study's goals and how they differ from other types of studies with which forensic scientists might be familiar). In addition, social and physical scientists use different scientific terminology. During the course of this project, much translation was needed by both sides in working with partners. The heavily technical nature of the grant activities also required substantial research and explanation, even for the team's forensic experts, as some of the interventions were so novel that they had little history in the field.

The research team was fortunate to have both social science and forensic science experts working together to produce research that would be useful for both groups. The research team also attempted to make the present document user-friendly for both audiences. We have paired intervention descriptions with more simple explanations directly in the text, supplemented by footnotes that provide more technical information on forensic issues and methodological or statistical issues that provide more basic information for those less familiar with the forensic field.

Finally, as with any research project that partners with practitioners, the research team encountered some challenges related to fitting the additional responsibilities of a research study into practitioner routines that were already overburdened and overstretched. We were fortunate to have laboratory partners that were supportive and interested in the current study and made extensive efforts to provide the information and laboratory data requested by the project team. However, quite understandably, there were frequently delays in these tasks due

to a whole host of other day-to-day demands and the need to balance evaluation needs with the primary mission of the labs.

2.5.2 Data Challenges

In addition to the data deficiencies noted previously, there were challenges associated with the availability, acquisition, and quality of data. First, the data accessible through electronic LIMS databases were limited in scope. In particular, the completion dates of some stages of DNA processing were not recorded electronically, and no site stored *time* completion information.[27] Because no time information was available to the researchers, any reduction in turnaround time smaller than one day will not be seen. It is possible that smaller productivity and efficiency gains based upon smaller time intervals were not captured in the current study due to these data limitations.

An additional problem was the lack of data for DNA cases at Allegheny County due to the small number of cases completed at this lab (monthly counts or averages could be based on fewer than five cases for some months). There were larger numbers for the serology unit (which was another target of the grant), but analyses of the DNA processing may have lacked sufficient power to determine effects. In a way, it is fortunate that this issue happened with this particular lab, as its implementation occurred too late in the study period to examine grant outcomes anyway.

Another limitation of the study is that there were no measures of quality or accuracy in DNA processing. Only one site had consistent information about reruns, and these occurred for such a small percentage of cases that it was unable to be analyzed separately. Other sites had this information in other locations (such as case files) but were unable to provide it in an efficient manner to the research team. Other measures of quality such as the outcomes of control samples in a batch were also not available to the research team. While we do not have separate quality measures, it is likely that quality problems will be reflected in longer turnaround times as samples needed to be rerun. Therefore, turnaround time should still provide a reliable measure of overall improvement in the lab process.

Other case characteristics were used to partition out the influence of these factors on turnaround time (e.g., priority level, analyst experience) so that the effects of the grant could be better isolated. However, electronic data also did not contain the individual-level information about samples/cases/routings that would have been helpful for analyses.

There were also resource indicator data limitations. Sites were only able to provide budget expenditures by year, if at all. At some sites, budget estimates also did not include

[27] The new LIMS for Kansas City recorded times; however, this was only available for nine months and was unreliable as the time recorded was the time of data entry and not necessarily the time the stage was completed.

important grant funds because sites could not account for such information by year. The method of counting lab staff also varied across sites. Some labs were able to provide start and end dates of staff employment, which allowed the team to better estimate the number of staff across the time period. However, others could only provide annual counts or the number of staff on a particular date each year (like a census). Annual resource denominators meant the efficiency estimates could only be reported by year.

Another data challenge was the varying units of analysis across sites. Two sites had electronic data tracked only at the case level, one site primarily tracked by sample (although some information was only case-level), and one site had enough data to track by either routing or case. As discussed previously, sites that record information at a unit of analysis higher than the level of processing will not have as strong an ability to detect changes. For example, if a laboratory analyzes samples individually (or by batches where the batch does not contain all case samples), but records data in the LIMS at the case-level (i.e., records the data when all samples for the case have undergone amplification), the data will not reflect the true date of completion for samples. The analyses will be insensitive to changes that impact the productivity and efficiency of samples but not cases overall.

Because available data depended on each individual site, the measures used at each site also had to vary. Differences across measures (e.g., start and end points defining turnaround time, whether canceled cases can be excluded from throughput measures) are discussed in greater detail above. However, it is important to note that these variations disallowed any comparisons across sites. We would also like to caution the reader more generally to not compare across the grantees, as their goals, scope of work, level of implementation, and data are too diverse to make analogous comparisons.

Another limitation of the structure of data was that the measures at the beginning and end of the study period have a higher potential for bias. Because data were extracted from January 2007 forward, the throughput measures at the beginning of the study period were likely underestimated, because they do not include cases begun before January 2007. There is also a likelihood that unfinished cases were concentrated at the end of the study period. Excluding these cases would only leave cases in the more recent months that tended to have shorter turnaround times. This would obviously bias the results towards finding grant effects that may not actually be present. In order to account for this, the research team used median imputation to estimate the expected turnaround time of unfinished cases based on the time spent on tasks completed thus far and the median time (based on all other completed cases) spent on incomplete tasks (see appendix B for more information). A final note on measures should be made that throughput measures are constrained by the amount of evidence submitted to a lab. In at least one site (UNT), staff reported that they often had to wait on additional submissions, to reach a minimum number of samples, before they could start batch-processing samples.

There were also issues with quality of laboratory data recorded. As discussed earlier, one site reported inconsistent data tracking practices. Most sites did not track important case information, such as new submissions or the need for reruns) consistently and clearly enough to be able to disentangle this. In the best situations, we were able to produce a dummy variable that identified cases which had one or the other of these; at other sites, we had no information on this. As long as these inconsistencies were equally likely in both the pre- and post-intervention periods, the analyses should still remain valid.

Finally, the data were quite "messy." There were an unexpectedly large number of outliers, nonsensical dates (e.g., October 12, 2032), chronologically out-of-order dates, and calculated turnaround times that were negative (due to the out-of-order dates) (see appendix B for more detail). While it is likely that some of these data problems are due to simple data entry errors, the research team also believes that much of the observed "problems" actually reflect pieces of reality. A variety of unique situations could explain some data discrepancies. For instance, turnaround times of over 20 years may be due to a data entry error or may be due to a "cold case" that is suddenly reexamined. New submissions of evidence may result in updating some, but not all, dates from previous sample processing. A sample may need to be reanalyzed years after its initial analysis in preparation for a trial. A report may be updated (and the date changed) after a match is made in CODIS, but no other processing dates are affected. New submissions may create a new assignment date but not influence other dates listed for stages that have not yet been completed. An analyst may enter today's date for a task they finished the previous week. A series of technical reviews for cases analyzed months or years earlier may be completed all at once in a single effort. Unfortunately, there was no way to determine with confidence which of the identified data issues were due to true, unique circumstances versus incorrect data entry. While some of these messy data may, in fact, be accurate, they nonetheless complicated the analyses and do not reflect the typical casework conducted at a lab. Therefore, these issues were addressed as described in appendix B to permit valid analyses.

2.5.3 Analysis Limitations

The largest limitation of the analyses is associated with the research design adopted for this exploratory evaluation. The research team relied on within-site pre/post and longitudinal assessments. Unfortunately, given the financial limitations of the contract award, comparison or control laboratories were not part of the study. As a result, analytic comparisons of grant-funded sites to others that did not receive any benefits of the grant program could not be made.

In addition, as was mentioned earlier, the post-implementation follow-up period of the study was constrained due to site implementation delays and contract limitations on the evaluation's period of performance. A substantial amount of baseline data was available;

however, at all of the sites, these baseline data were unstable and did not provide a clear picture of "business as usual" at the sites. Perhaps, a better interpretation is that "business as usual" for labs is rarely typical. Regardless, the comparison tests (t-tests and Mann-Whitney U tests) comparing pre-intervention monthly data to post-intervention monthly data have limited power to detect change due to the limited number of months in the study period and to the small sample size of post-intervention months for some sites (Allegheny only had 1–2 months of "post" data so no t-test was attempted, while Louisiana had only six months of "post" data).

The regression analyses are more powerful than the t-tests because the unit of analysis is the individual case/sample/routing, instead of monthly throughput numbers for the comparison analyses. In addition, the regression analyses controlled for other factors expected to influence turnaround time. However, one notable limitation of regression analysis is that it does not prove causation. These analyses can detect relationships between independent or "predictive" variables and the dependent or outcome variable. There is no way to determine with confidence if an observed relationship is due to the independent variable *causing* changes in the outcome as opposed to (a) the outcome variable causing changes in the independent variable (although this can often be ruled out based on common sense and chronological ordering of events) or (b) due to a third confounding variable. We can only speculate on the potential causality of an observed effect based on what we know through other means. Another limitation of the regression analyses regards the inclusion of variables in the models. The absence of important variables that influence turnaround time may result in biased estimates since the effects of these additional variables cannot be accounted for. As mentioned earlier, the team did not acquire all the information requested because much of this data were not recorded electronically. Further, some variables (such as priority status or offense type) had to be excluded from regression analyses because the high percentage of missing values eroded the sample size significantly.

In addition, a proxy variable was used to control for individual analyst performance. It was determined that the most objective measure would be the years of professional experience. However, this is not a perfect measure. In addition to the fact that capabilities and quality of work is not always related to experience, there are also potential selection biases with this measure that may counteract the expected effect. For instance, senior staff may be given more difficult cases that might take longer to analyze or senior staff may take longer to process a case when they are juggling other responsibilities such as supervision or management. Batch and team processing may also make "ownership" of a case difficult to determine and dilute the effects of analyst experience.

Another challenge of the regression control variables was that the research team had to find a way to control for different types of evidence items. In categorizing items, the team focused on items that might require additional time spent on "searching," such as textile

items; samples that might require differential extractions or mixtures, such as intimate samples; and more simple items, such as swabs and known standards that require less screening work.

Finally, regression analyses could not be used to analyze throughput measures because throughput is a monthly count and not measured at the case/sample/routing-level. Also, as noted previously, throughput measures should be interpreted with caution at the beginning and end of the study period. Cases begun before January 2007 and cases begun toward the end of the study period that were completed after January 2011 were not included in the data. Therefore, the counts for the first and last few months of the study period may be underestimated.

2.5.4 Limitations of the UNT Timed Experiment

There are several limitations to the UNT timed experiment design that should be considered when interpreting any results from this supplementary analysis. First, the two divisions collected time data differently. The analyst in the research division recorded all date and time data as the DNA processing steps were being performed.[28] The casework division recorded time data for batches already processed. The casework time data were reported as the total minutes worked, as opposed to actual start and stop times. These data were retrieved from casework division laboratory documentation and from an analyst's recollection of how long specific steps took. As a result, each stage was treated independently, and the determination of the total DNA processing time for each method was not possible. However, the total time difference for the individual substages was calculated.

The second limitation of this timed experiment was the low number of observations for each DNA processing stage. Each division provided time data for three batches.[29] While these batches consisted of many samples, which developed hundreds of STR alleles and thousands of mtDNA base calls, the actual DNA processing was performed only 3 to 11 times.[30] Because the number of observations is so low, the average time differences between the two methods lack the power necessary to draw conclusions on whether the differences are large enough to be considered statistically different.

The third limitation of this timed experiment was the comparability of the data supplied by both divisions. While there were three sets of batches being compared, only one set

[28] This real-time data collection was the intent of the timed experiment and would have allowed for comparisons both at the stage level and for the process as a whole.

[29] And, the research division supplied additional stage times from partially completed batches and tests of the new dye system.

[30] Substages could have more than three observations due to repeated actions within the same batch (reinjections, reamplifications), which were treated as separate observations of that substage.

included the same samples being processed by both methods. The other two sets processed by each division contained different family reference samples. As a result, any stage that would be affected by sample quality or quantity could not be directly compared. In order to increase the comparability of these data sets, any repeat of a substage within a single batch was treated as a separate time observation for that substage. This was done because the number of repeated steps (e.g., reinjection, reamplifications) is a function of sample quality and quantity and the DNA processing procedures. The research division provided additional DNA processing quality data, that compared the results of the same three batches processed with both methods. These self-reported results are presented separately from the timed experiment in section 7.4.4.

3. CASE STUDY: ALLEGHENY COUNTY MEDICAL EXAMINER'S OFFICE FORENSIC LABORATORY DIVISION

3.1 Overview of the Laboratory

The Allegheny County Medical Examiner's Office Forensic Laboratory Division (hereafter, Allegheny County) is an accredited public crime laboratory housed within the Allegheny County Police Department in Pittsburgh, Pennsylvania. The laboratory accepts DNA samples from 137 agencies in the county.

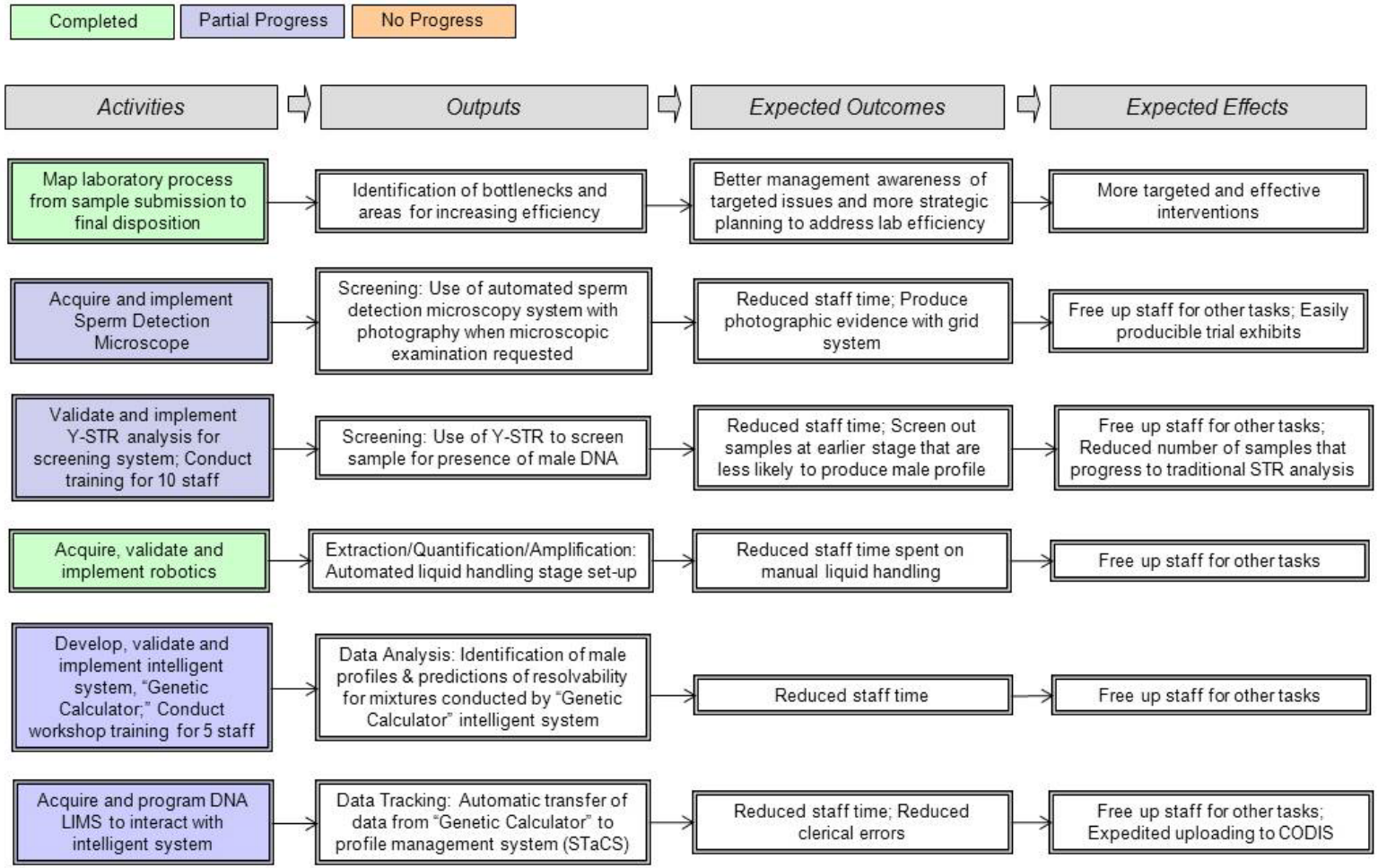
Allegheny County's serology and DNA units work together to analyze evidence from homicide, sexual assault, robbery, and burglary cases. The serology unit examines submitted items for stains from physiological fluids that may yield probative DNA profiles. For sexual assault cases, serologists conduct manual microscopic examinations of evidence samples to identify the presence of sperm. After serological examination, evidence is transferred to the DNA analysts, who begin the process of developing a DNA profile. Manual microscopic screening for sperm can be very time-intensive (taking, on average, 16 hours) and was identified as a bottleneck in the beginning stages of DNA processing.

Allegheny County proposed to identify other bottlenecks in the sexual assault evidence workflow through process mapping in order to improve efficiency for the handling and analysis of sexual assault evidence. The tentatively planned approach (dependent on process mapping results) included a new automated sperm search microscope, utilization of Y-STR analysis as a screening tool for male DNA in mixture samples, the use of an intelligent "genetic calculator" expert system, automatic transfer of data from the expert system to a DNA laboratory information management system module, and the minimization of manual, repetitive liquidhandling tasks through the implementation of robotic systems.

3.2 Description of Grant Goals

NIJ awarded Allegheny County a $382,309 grant to fund personnel costs, purchase of an automated sperm-detection microscope and automated liquid-handling robotic workstation, development of an expert system, and maintenance costs. The local 25 percent nonfederal match supported the expansion of their LIMS, purchase of the expert system processor unit, and training costs. Allegheny County described five main goals of the proposed strategy to improve efficiency within their lab (see figure 3). These goals, the activities involved in achieving these goals, and the expected outcomes and effects are described below. The lab's proposal predicted that the funded changes would create an 80 percent decrease in processing time.

Figure 3. Logic Model of Proposed Interventions for Allegheny County



Allegheny County proposed to partner with a LIMS provider to map their laboratory process and identify bottlenecks and areas for increased efficiency. By addressing identified bottlenecks, it was assumed that efficiency could be markedly improved. Further, if efficiency could be increased and the backlog reduced, the lab expected to be able to serve a more investigative function in casework, as opposed to primarily analyzing evidence in cases with set trial dates, per their prioritization practices.

Allegheny County's proposed plan, contingent on the results from the process mapping, was to improve screening of sexual assault evidence and create an expert system integrated with the lab's LIMS. In order to increase efficiency at the screening stage, they planned to validate the use of Y-STR analysis as a method of screening sample submissions. Rather than manually examining all samples for the presence of sperm and then proceeding with analysis, DNA would first be extracted from the sample, and then they would use Y-STR analysis to determine whether DNA from a male contributor was present. They then proposed to introduce the *Cybergenetics TrueAllele* software, an artificial intelligence expert system that identifies alleles from an electropherogram (the data output of capillary electrophoresis). During the screening stages, the *TrueAllele* system would be used to examine data to determine whether DNA from a male contributor is present and to predict the resolvability of a mixed sample with multiple male DNA contributors. Allegheny's proposal also included training for five analysts on the *TrueAllele* system.

Using Y-STR analysis as a screening tool was expected to reduce the need for manual serological examination and decrease the number of samples that would progress to traditional, autosomal STR analysis (i.e., a profile from multiple chromosomes, not just the Y chromosome). Because the serological screening work was very time consuming, this new Y-STR screening process was expected to lead to a reduction in staff time, which would free up staff for other lab tasks.

For cases where a microscopic identification of spermatozoa was requested, or would be a preferred exhibit in court, Allegheny County planned to use an automated sperm detection microscope with a novel sorting feature (images sorted by likelihood of having sperm-like morphological features) and the ability to photograph the samples for use as evidence. Training for 10 analysts was included in the proposal budget. The novel sorting feature was expected to reduce staff time needed for direct examination, while photographic evidence would assist with court presentations.

Once the evidence items were screened through the above-listed methods, they would then advance to traditional batch processing and undergo traditional DNA typing to produce autosomal STR profiles. Allegheny County proposed to use robotics platforms to institute automated liquid handling at multiple stages in the DNA analysis process. It was expected that two *Biomek 3000* robots would be validated and implemented at the extraction and

quantification steps while a *JANUS* robot, acquired through this grant, would be used at the amplification stage. The utilization of robotic systems was expected to replace the need for staff to transfer liquids, which would then free up staff for other lab tasks.

Once DNA typing was complete, Allegheny County proposed to use the *TrueAllele* expert system on the back end to identify alleles for the STR profiles obtained from autosomal analysis. The *TrueAllele* system was supposed to produce a match-likelihood ratio for comparison of DNA profiles from questioned evidence[31] to profiles from reference DNA samples.[32] Rather than have two analysts independently interpret the data, the *TrueAllele* expert system could perform the first data review and interpretation,[33] and only one analyst would be needed for further data review. As a result, the use of this expert system was expected to reduce staff time spent on the analysis stage of DNA processing.

The final component of Allegheny County's efficiency improvement strategy was to integrate the new expert system with a DNA profile management system so data were automatically transferred from the *TrueAllele* software to the new profile management system, *STaCS*. In order to do this, Allegheny County needed to select a vendor through the county's competitive bid process; obtain and install the server, hardware, and back-up software; validate the software; and program an interface for the system to communicate with *TrueAllele*. The automatic transfer of these data was expected to lead to a reduction in staff time spent on manually moving the data and a reduction in clerical errors.

[31] Evidence where the source of the DNA profile is unknown.

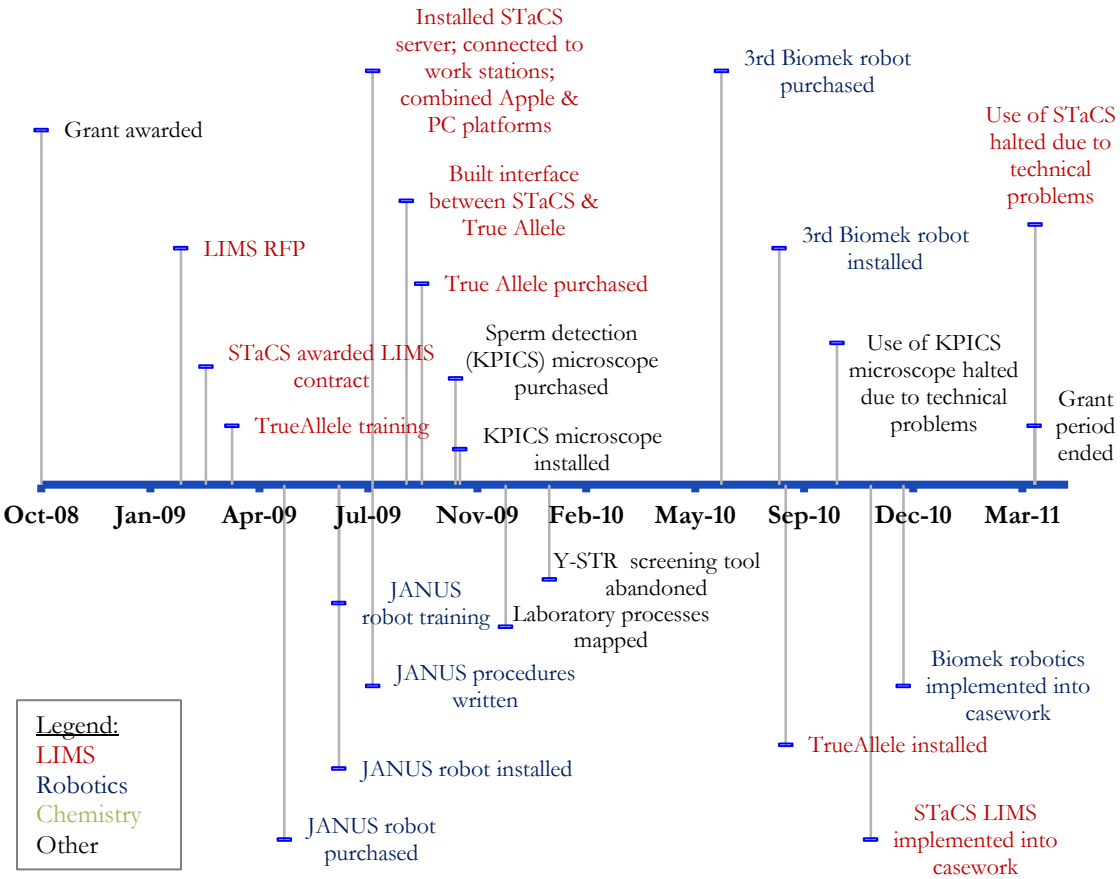
[32] Samples collected directly from a person, hence the source of the DNA profile is known.

[33] Because "analysis" is used frequently in this report to reference the current evaluation's analyses, the word "interpretation" will be used for the DNA analysis stage to avoid confusion between the two concepts.

3.3 Implementation Findings

3.3.1 Implementation Description

Figure 4. Allegheny Implementation Timeline



Allegheny County concluded their grant activities on March 31, 2011, 29 months after notification of their award on the grant activities. They received multiple extensions to support a number of implementation challenges, which are discussed below.

The lab began by process mapping their serology and DNA unit activities through coordination with the Division of Computer Services, *Cybergenetics*, *STaCS DNA*, and laboratory staff. The following bottlenecks were identified in the process mapping report:

(a) Manual serology screening: DNA analysts were cutting samples from larger stains removed from garments during serological analysis. This required significant manipulation during extraction setup. This was not only time consuming but also

introduced a potential point for contamination between samples because of the number of samples being manipulated, according to the report.

(b) Manual semen screening: Manual microscopic screening of suspected semen specimens was very time consuming.

(c) Data entry: Entry of data into DNA LIMS (DLIMS) at the time of extraction was an inefficient use of time and redundant processing.

(d) Lack of redundancy: Dedicated liquid handling instrumentation for extraction and quantification was a potential bottleneck. Dedicated systems were more efficient, provided consistent functioning. However, if one of the components malfunctioned, the lack of redundancy would result in the entire system shutting down.

(e) Computer operating systems: The DNA LIMS system was a Windows-based application, while the *TrueAllele* expert system was Mac based, making reporting and review a cumbersome task.

After conducting the process mapping, Allegheny County decided to move forward with the proposed plan. The lab purchased the *Niche Vision KPICS Sperm Finder* system. This technology uses the Kernechtrot-Picroindigocarmine (KPIC) stain to visualize sperm and a sorting algorithm to automatically detect sperm by their morphological characteristics. During the grant period, the instrument was installed and one lab analyst received training on this system. It was necessary to install an anti-vibration pad to stabilize the system and allow for more accurate image focusing. Validation was completed on the system by mid-2010. However, by the end of the grant period this technology was not being used in casework due to ongoing IT and technological issues.

After a competitive request for proposals in February 2009, Allegheny County selected *STaCS DNA*, Inc., as the vendor for the development of the profile management system and integration of this system with the *TrueAllele* expert system. However, due to incompatible *STaCS* LIMS (PC) and the *TrueAllele* system (Mac) platforms, the project director had to locate and use *VMFusion* software in order to create a computer environment that was compatible to both the Macintosh and PC platforms. By July 2010, Allegheny was using the *STaCS* system to track calibration, instrument maintenance, and usage of supplies and reagents. Allegheny planned to extend the *STaCS* system to track samples after the robotics and sperm detection microscope were fully implemented. By the end of the project, technical and other IT issues hampered complete implementation of the system. Laboratory project managers no longer intended to utilize the *STaCS* system as originally planned and were shifting to customize the *PorterLee BEAST* LIMS to meet their needs.

The *TrueAllele* system was purchased and interfaced with *STaCS* in September 2009. Training was conducted for one analyst on the use of the *TrueAllele* system with the

traditional manifold. The lab worked toward validating the use of *TrueAllele* for the more traditional data-interpretation purpose on the back end but did not complete this validation by the end of the study. No progress was made during the grant period toward the use of *TrueAllele* for the nontraditional purpose of male DNA screening (see section 3.3.2, Implementation Challenges, for more information about this).

The originally proposed workflow included one *JANUS* robot, along with two *Biomek 3000* robots for liquid handling. The *JANUS* robot was purchased in May 2009 and installed by July of that year. (Three *Biomek 3000* robotic systems were purchased with funds outside of this grant.) One member of the laboratory staff was trained on the *JANUS* instrument in July. However, after the loss of knowledge through personnel changes, plans to validate and implement the *JANUS* robot were put on hold. While this technology has more advanced capabilities than other robotics, the effort to learn two new robotics systems was deemed not worth the capabilities gained. Lab staff instead shifted toward making all DNA processing steps (extraction, quantification and amplification) automated with the *Biomek* systems. By the end of the grant period, Allegheny had installed, validated, and implemented two of the three *Biomek* robotic platforms in casework, and the third was implemented after the grant period ended. On the robotic platforms, the *DNA IQ* extraction kit replaced manual organic extraction, and the *Plexor HY* quantitation kit replaced the *Quantifiler* kit used in the manual procedure.

The performance of Applied Biosystem's *Yfiler* amplification kit was tested and compared to Promega's *Powerplex-Y* system. After this comparison, Allegheny decided to proceed with the *Yfiler* kit. Validation work on the *Yfiler* system was completed during the grant period; however, this kit was not implemented into DNA unit processes until the end of the grant. Allegheny reported that additional testing was needed to develop interpretational guidelines, and, by the end of the study, staff training was ongoing. Once these tasks are completed, Allegheny anticipates that the *Yfiler* kit will be fully implemented. However, no Y-STR analysis will be used as a screening tool, in lieu of traditional serology screening, as originally proposed by the site.

By the end of the grant period, Allegheny County had implemented some of the new chemistries and robotics described above. They also continued to maintain all materials and procedures for the manual processes since some staff were still performing tasks manually while they awaited training. Additionally, if the automated systems ever go out of service, lab staff can resume manual procedures until they are back online.

3.3.2 Implementation Challenges

While Allegheny County began the development of their grant activities in an expeditious manner, they encountered a number of challenges that delayed full implementation of the grant goals. The first delay was caused by a move to a new laboratory. While many of the

grant components had been individually developed and validated in their original lab, it was decided to wait on purchasing the sperm-detection microscope and validating the grant components as a complete and integrated system until they moved into the new laboratory. However, the date of this move was pushed back multiple times and did not occur until July 2009. DNA casework was not reinstituted until October 2009 (serology casework began two weeks after the move was completed).

There were further difficulties in developing the expert system for the Y-STR and mixture analysis. The algorithms were more complicated than the vendor originally anticipated, leading to additional delays. As a result of this delay and some changes to the project staffing, the lab reevaluated the usefulness of Y-STR analysis to replace serological screening and decided to abandon plans to implement this male DNA screening system. In their final report, project staff stated that even if this technology was implemented as a male DNA screening tool, significant information about the source of the male DNA would be lost. While this technology could elucidate DNA from male contributors, it could not determine the physiological source of that DNA (i.e., saliva, semen, blood etc.). The source fluid could be important contextual information depending on the investigation.

An additional obstacle to implementation occurred in January 2010, when a change in project leadership occurred. The original project manager had been developing the grant program together with one other analyst. This project manager and analyst left the lab, and grant management duties were transferred to another analyst. However, a loss of information occurred with this transfer. No remaining lab staff had substantial knowledge about the grant goals and activities. This caused a significant delay in grant progress, as the new project manager and other lab staff needed to take time to familiarize themselves with the project and its status. The original project manager also reported a slow county acquisition process as an additional implementation challenge.

3.3.3 Final Perceptions

At the end of the study period, the key contact at Allegheny County reflected on the perceived impacts of the grant on the lab, lessons learned, and future plans. The largest benefits of the program were viewed as the available funding to develop new approaches and the lab's increased throughput capability from robotics. The lab said they would not have had the funds to purchase new equipment or supplies or to test the new procedures without the grant. Allegheny reported that their records showed they had issued 157 reports as of September 2011, compared with 91 reports in all of 2010. If one projects the 2011 figure out to a full calendar year, the lab might expect around 209 cases completed in 2011, 2.3 times the number of reports in 2010. The lab attributes these gains to the implemented robotics for extraction, quantification, and amplification. Further, the lab felt that these changes helped

their clients obtain information more quickly and possibly helped them to identify and remove offenders from communities faster.

The biggest lesson learned was the danger of isolating this project to a small number of staff. Instead, the interviewee reported that the lab would try to build larger teams and delegate more in order to avoid drastic losses of institutional knowledge for future projects. The point of contact also said they would prefer to implement pieces more slowly instead of dramatically changing the workflow in multiple ways simultaneously.

Allegheny reported high satisfaction with NIJ and the grant program, commending the clear expectations and networking opportunities (e.g., NIJ Conference). In particular, the interviewee thought the competitive structure and emphasis on *efficiency* motivated labs to think about backlog problems in new ways and develop novel, exciting solutions. Although the lab was not able to achieve all of its goals at the time of this report, it is continuing to work on validating the *TrueAllele* expert system, communicating with the automated sperm-detection microscope vendor to fix technical problems, and replacing the DNA LIMS module.

3.4 Outcome Findings

The following section describes the data used to assess the outcomes of the NIJ Forensic DNA Unit Efficiency Improvement Program on Allegheny County and changes in productivity and efficiency at Allegheny County.

3.4.1 Descriptive Statistics and Trend Analysis

Allegheny County performed analysis work on 1,511 forensic serology cases during the evaluation period.[34] One-quarter (25.5 percent) of these serology cases proceeded onto DNA processing. The majority (77.2 percent) of cases were for violent offenses, with 23.8 percent involving sexual assault incidents (the original target of the grant). Homicides made up over one-third (36.5 percent) of cases, while there were small numbers of property (9.1 percent) and drug (0.6 percent) offenses. There was a reported suspect, at the time of DNA processing, in around half (55.5 percent) of all cases.

Serology cases had, on average, 6.25 items submitted (it is unknown how many of these items moved on to DNA analysis), including 38.2 percent of cases with a submitted sex assault kit and 25.0 percent of cases with some form of textile evidence (including clothing or bedding). At least one in five cases (21.4 percent) experienced a rerun or multiple submissions at some point during the case's history. There were 13 serologists and six DNA analysts reported to be responsible for these cases across the study period.

[34] 1/1/2007–1/31/2011.

Median monthly throughput outcome measures were 27 cases for serology and 5 cases for DNA across the study period (see table 3). The median turnaround times were 7 and 59 days, respectively, for serology and DNA casework. Before assignment (when the defined turnaround time began), there was a median of 75 days between evidence submission and assignment to the serology unit. A median 28 days passed between the completion of serology work (defined as the administrative review) and DNA extraction. Stages of DNA processing (including those between extraction and quantification, quantification and amplification, amplification and capillary electrophoresis, capillary electrophoresis and report completion, and report and administrative review) had median turnaround times of 1–23 days. The shortest turnaround time was between amplification and capillary electrophoresis, while the longest was between electrophoresis and completion of the report.

Statistics for efficiency indices (throughput and turnaround time divided by annual labor counts and budget expenditure estimates [in $100,000 units]) are also shown in table 3.

Table 3. Allegheny County Throughput and Turnaround Time Outcomes

Serology cases (*N* = 1,511), DNA cases (*N* = 365)		Productivity/Labor	Productivity/Budget	Cleaned Productivity	Raw Productivity
Overall Outcomes					
Serology Case Turnaround Time	Mean	2.26	33.36	21.35	44.78
(Assignment–Admin Review)	Median	0.72	11.33	7.00	8.00
	Std. Dev.	5.01	74.60	47.50	120.97
	Range	(0, 39.72)	(0, 636.80)	(0, 384)	(0, 1347)
DNA Case Turnaround Time	Mean	10.58	152.65	99.48	139.96
(Extraction–Admin Review)	Median	6.20	91.21	59.00	62.00
	Std. Dev.	12.51	181.43	121.14	314.75
	Range	(.90, 99.50)	(12.75, 1409.74)	(9, 995)	(-225, 2855)
Serology Case Throughput	Mean	3.24	46.01	N/A	30.29
	Median	2.79	41.65	N/A	27.00
	Std. Dev.	1.39	19.17	N/A	11.96
	Range	(0.83, 6.17)	(0.09, 0.43)	N/A	(9, 60)
DNA Case Throughput	Mean	0.63	9.14	N/A	5.80
	Median	0.59	8.18	N/A	5.00
	Std. Dev.	0.50	7.37	N/A	4.22
	Range	(0, 2.51)	(0, 0.17)	N/A	(0, 22)
Stage-Level Turnaround Time					
Serology Submission–Assignment	Mean	10.69	152.98	98.96	106.60
	Median	7.89	108.83	75.00	76.00
	Std. Dev.	10.07	146.92	92.50	128.01
	Range	(0, 57.39)	(0, 841.72)	(0, 483)	(0, 2267)

Stage-Level Turnaround Time		Productivity/Labor	Productivity/Budget	Cleaned Productivity	Raw Productivity
DNA Admin Review–Extraction	Mean	8.38	124.31	80.19	-63.51
	Median	3.00	43.57	28.00	8.00
	Std. Dev.	16.51	255.84	160.98	454.85
	Range	(0, 118.76)	(0, 1903.77)	(0, 1148)	(-4167, 1148)
Extraction–Quantification	Mean	1.16	16.84	10.87	-71.46
	Median	0.70	9.92	6.00	6.00
	Std. Dev.	2.08	30.25	19.47	852.99
	Range	(0, 17.38)	(0, 278.60)	(0, 168)	(-14464, 371)
Quantification–Amplification	Mean	0.55	8.06	5.11	56.05
	Median	0.31	4.98	3.00	3.00
	Std. Dev.	0.94	13.95	8.22	832.27
	Range	(0, 13.19)	(0, 193.44)	(0, 111)	(-363, 14611)
Amplification–CE Injection	Mean	0.28	3.99	2.59	-3.03
	Median	0.10	1.66	1.00	1.00
	Std. Dev.	0.45	6.47	4.28	67.41
	Range	(0, 3.25)	(0, 42.99)	(0, 32)	(-1093, 32)
CE Injection–Report	Mean	7.21	104.91	68.56	111.13
	Median	2.50	35.55	23.00	24.00
	Std. Dev.	14.91	219.80	145.09	346.92
	Range	(0, 116.75)	(0, 1820.86)	(0, 1148)	(-2176, 2725)
Report–Admin Review	Mean	3.10	44.47	28.85	48.32
	Median	1.07	15.68	10.00	10.00
	Std. Dev.	5.99	85.28	55.97	237.03
	Range	(0, 44.67)	(0, 655. 25)	(0, 391)	(-8, 3029)

Notes: Labor is defined as the number of staff reported for that year. Budget is defined as the annual DNA unit budget in $100,000 units. Turnaround time is reported in number of days.

There was substantial variability in these outcome measures across *cases*, as shown in the standard deviation and range statistics and as discussed in greater detail in appendix B. In addition, there was substantial variability across the *study period*. Figures 5–6 and 9–10 show both the number of completed and started cases[35] across the study period with vertical lines indicating the date of implementation milestones, including the implementation of a LIMS DNA module and robotics.

Turnaround time also varied by month (see figures 13 and 16). Some of this variation may be related to known events occurring in the lab. Implementation of the *STaCS* DNA LIMS module and robotics into casework occurred too late in the study period to determine with confidence whether there was any change in throughput or turnaround time due to these changes. Further, the case processing outcomes before this point did not produce a stable baseline due to high variability over time. Examining trends over time for turnaround time of individual *stages* of DNA processing also revealed a wide variation across months with no consistent or clear pattern detected (see figures 19–23).[36]

Other changes at the lab occurring earlier might be responsible for some of the variation. For instance, the graphs reveal a decrease in DNA throughput between August and November 2009, immediately after the organization's July move to a new lab. The lab stated that DNA casework was not fully functioning until October 2009, and the data reflect this. The lab reported that serology casework was only disrupted for the month of the move, which explains why less of a decline is shown in throughput for serology. In June and July 2010, there were interns working at the lab, and a quantification instrument was implemented through another grant. It is difficult to determine whether these changes had any impact on serology, which shows a slight decrease in throughput, but surrounding months are also quite variable. The DNA unit, however, appears to be completing fewer DNA cases compared to adjacent months. While the quantification instrument might be expected to improve productivity, the presence of interns could have slowed down casework due to the need for intensive training and supervision.

When taking into account staff and budgetary resources, 2010 appears to be the most efficient year for both serology and DNA throughput (see figures 7–8 and 11–12). For serology cases, at least, the lab has been continuously improving its efficiency each year. Efficiency measures of serology turnaround time (figures 14–15) reveal lower efficiency in 2007 compared with the fairly stable efficiency estimates for 2008–10 (greater efficiency is

[35] Completed cases are not necessarily the same cases as those started each month. Started cases are matched to the month in which a case was assigned, while completed cases are assigned to the month in which the case was completed.

[36] Only productivity measures are shown for stage-level turnaround time because the late implementation does not allow for analysis of grant impacts and the efficiency denominators did not substantially change trends found for overall productivity measures.

indicated by higher estimates for throughput and lower estimates for turnaround time). DNA casework, on the other hand, appeared to be more efficient with staff and budgetary resources in 2007 and 2008 (figures 17–18). Again, any changes evident in the data are not due to the grant program, as implementation did not occur until the end of the study period.

Figure 5. Monthly Number of Serology Cases Started

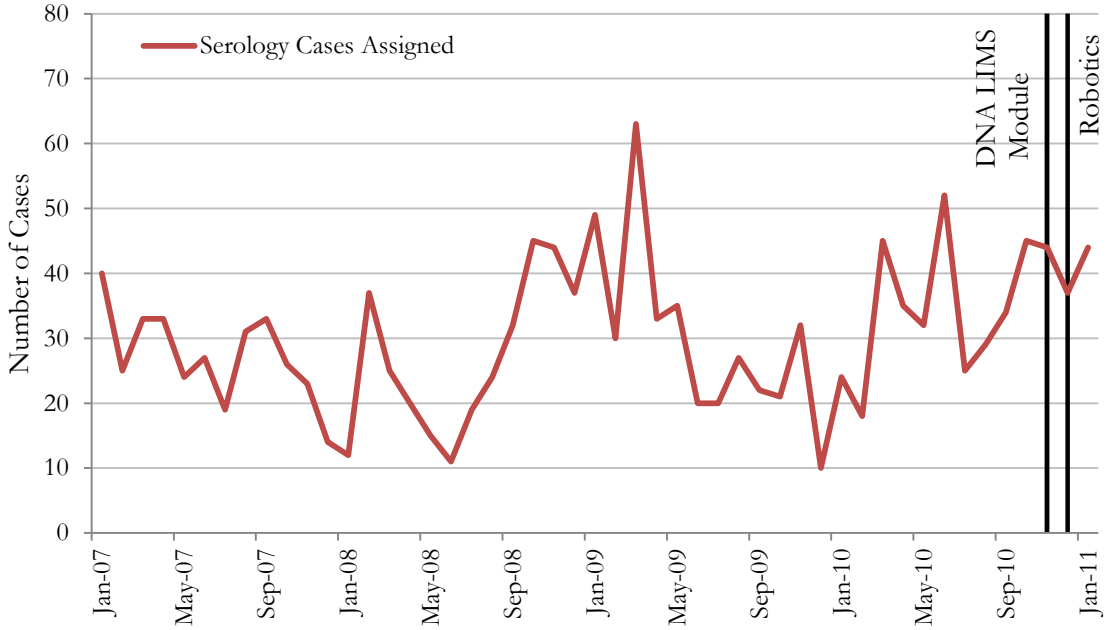


Figure 6. Monthly Throughput of Serology Cases

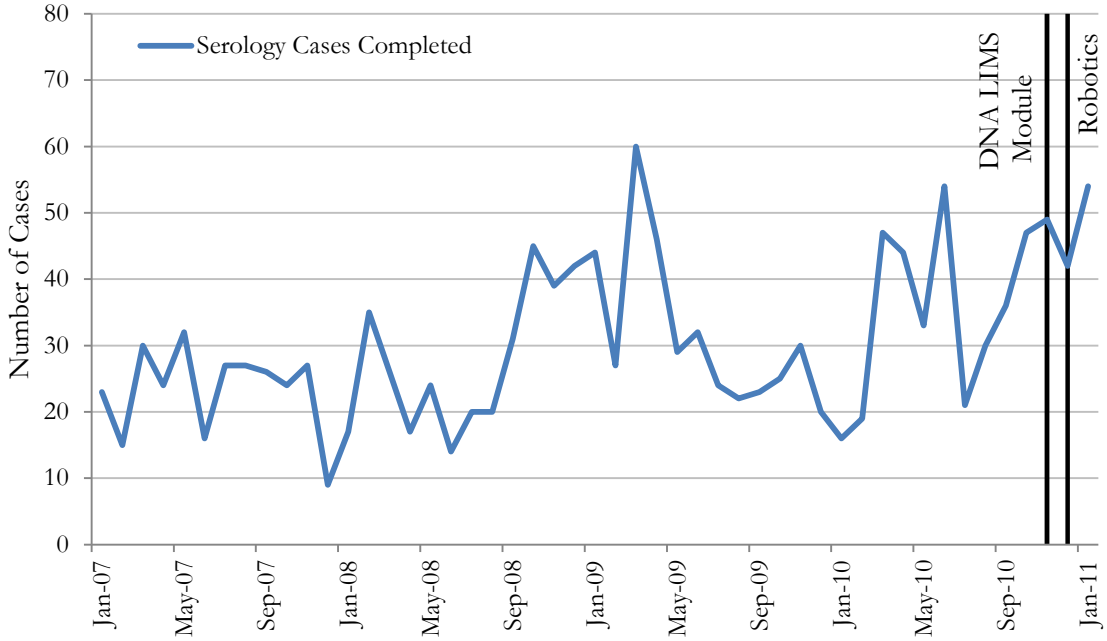


Figure 7. Efficiency Measure of Monthly Serology Throughput by Labor

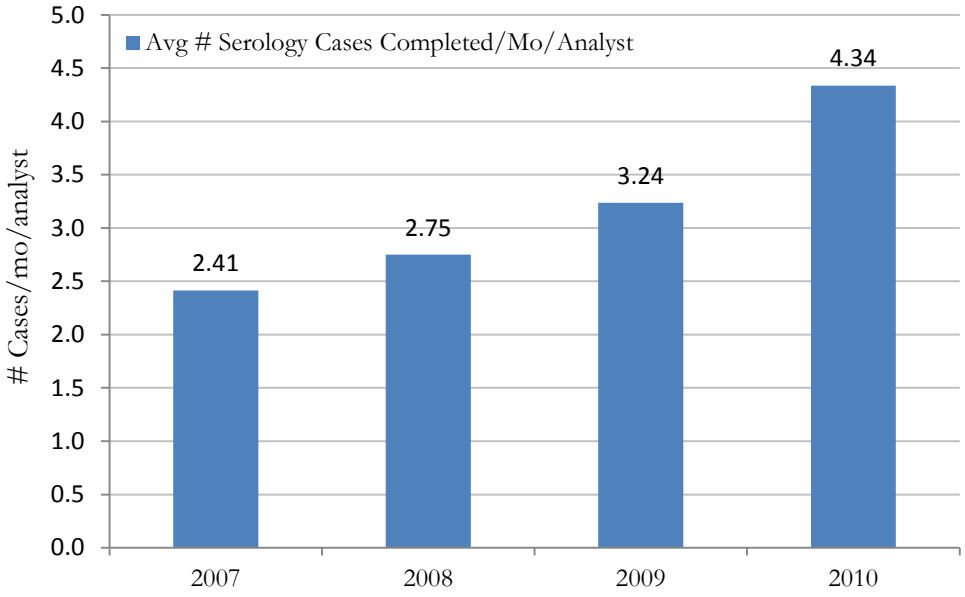


Figure 8. Efficiency Measure of Monthly Serology Throughput by Budget Expenditures

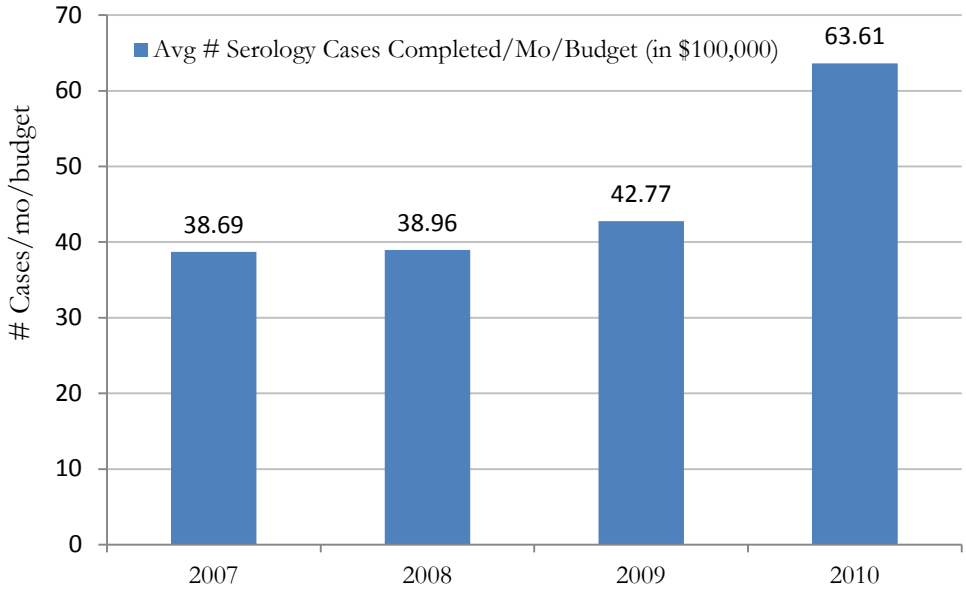


Figure 9. Monthly Number of DNA Cases Started

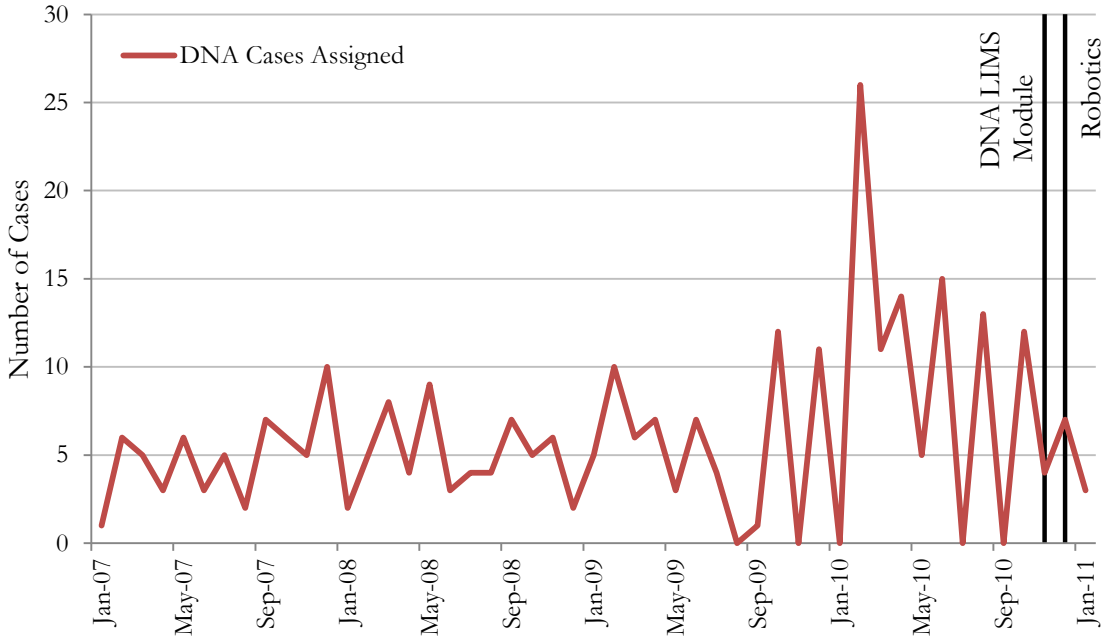


Figure 10. Monthly Throughput of DNA Cases

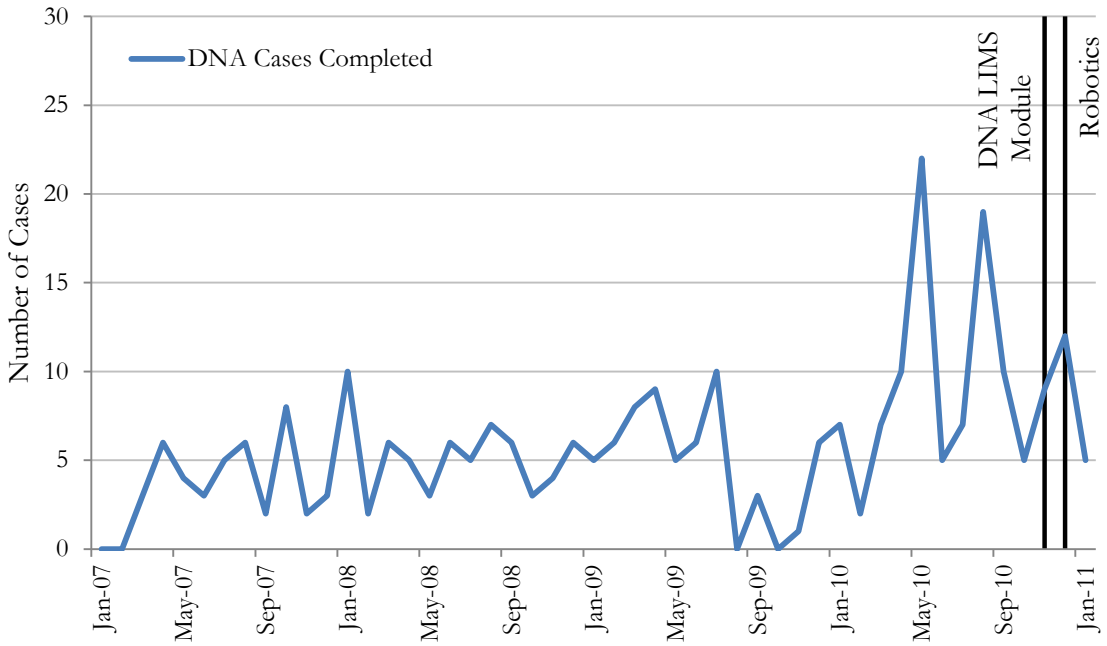


Figure 11. Efficiency Measure of Monthly DNA Throughput by Labor

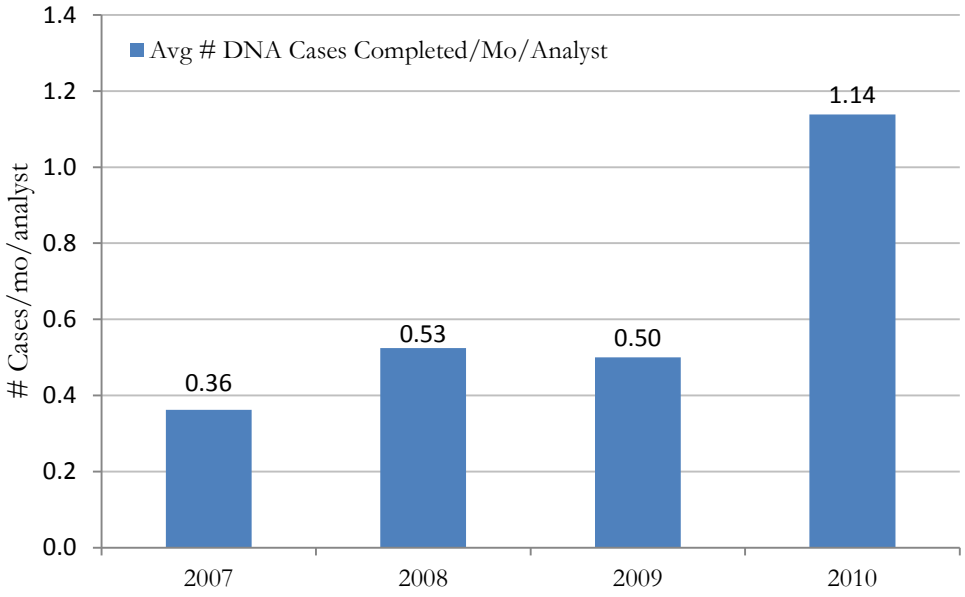


Figure 12. Efficiency Measure of Monthly DNA Throughput by Budget Expenditures

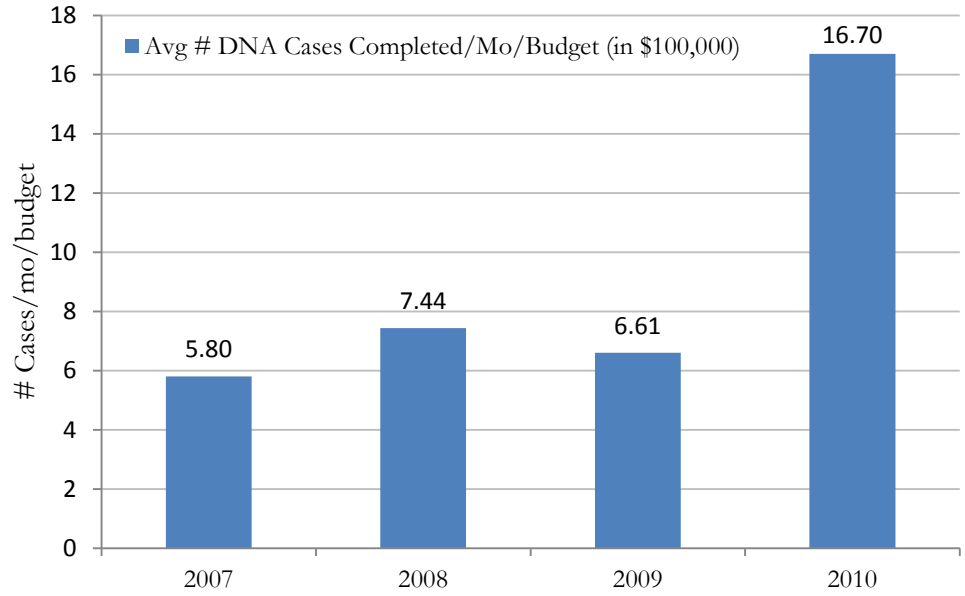


Figure 13. Serology Case Turnaround Time

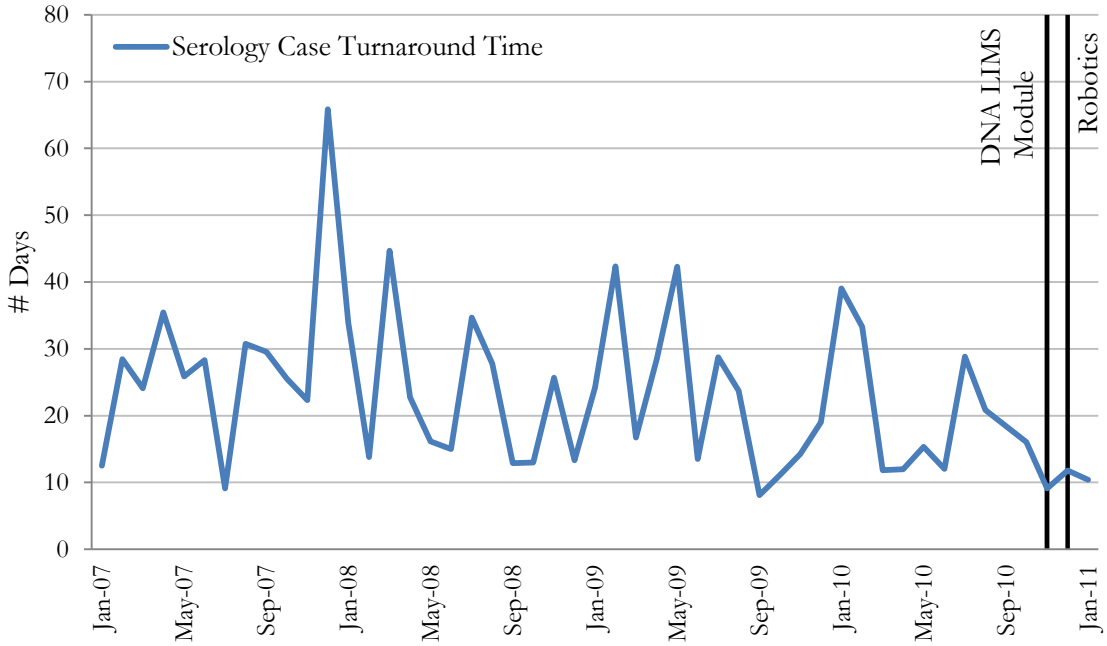


Figure 14. Efficiency Measure of Serology Turnaround Time by Labor

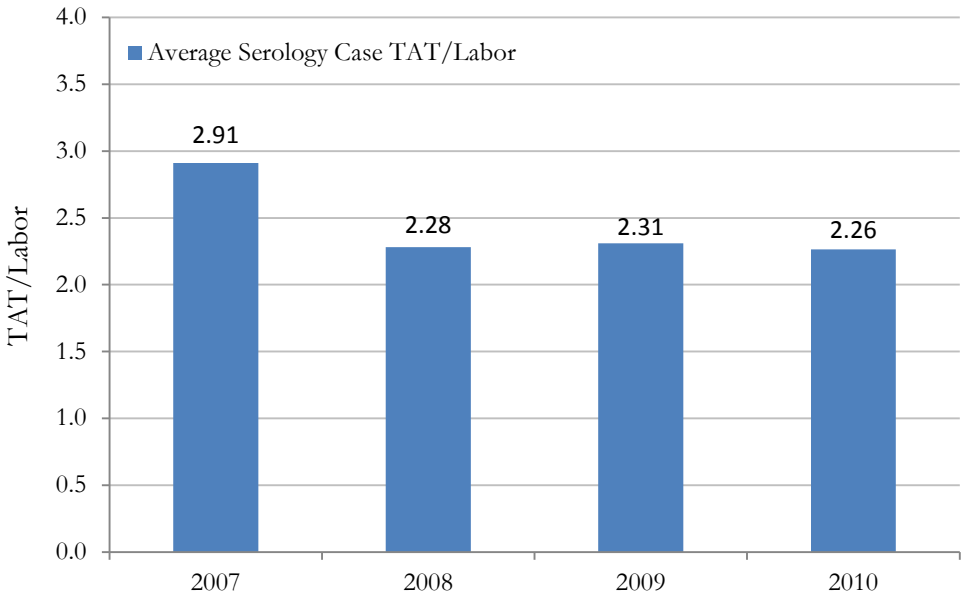


Figure 15. Efficiency Measure of Serology Turnaround Time by Budget Expenditures

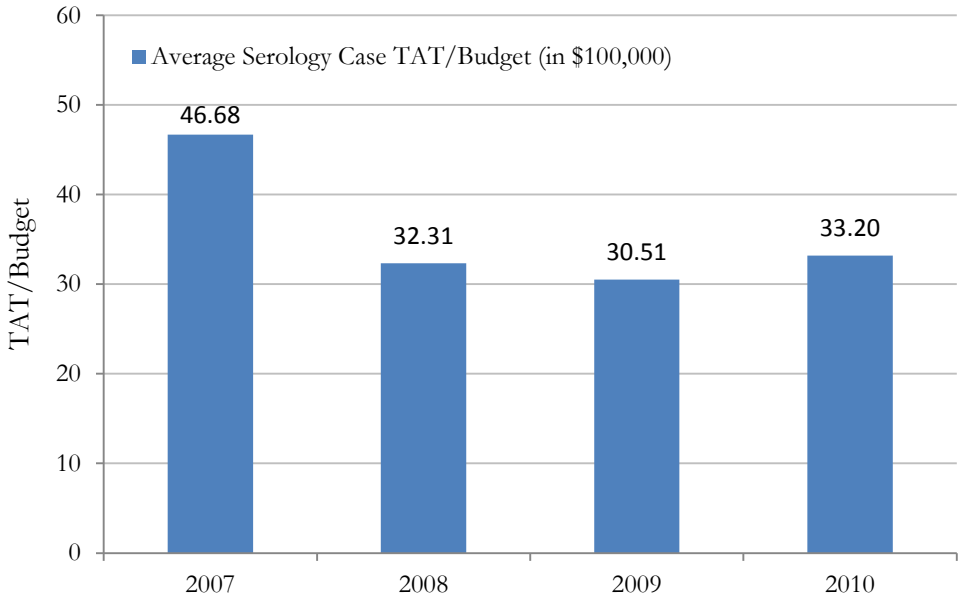
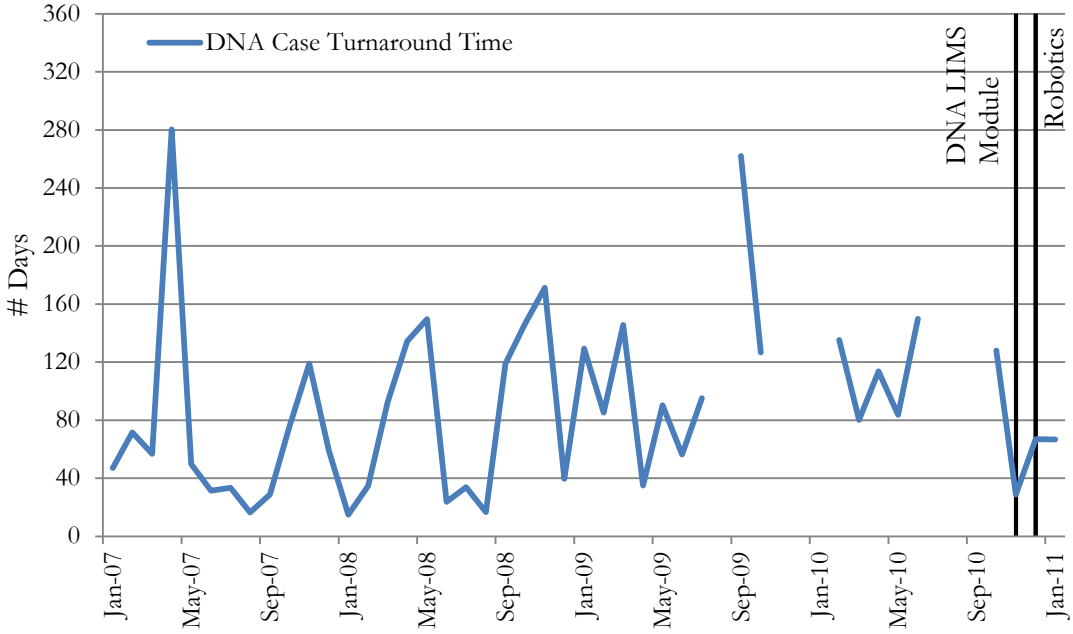


Figure 16. DNA Case Turnaround Time



Note: Months with missing data did not have any DNA cases assigned that month.

Figure 17. Efficiency Measure of DNA Turnaround Time by Labor

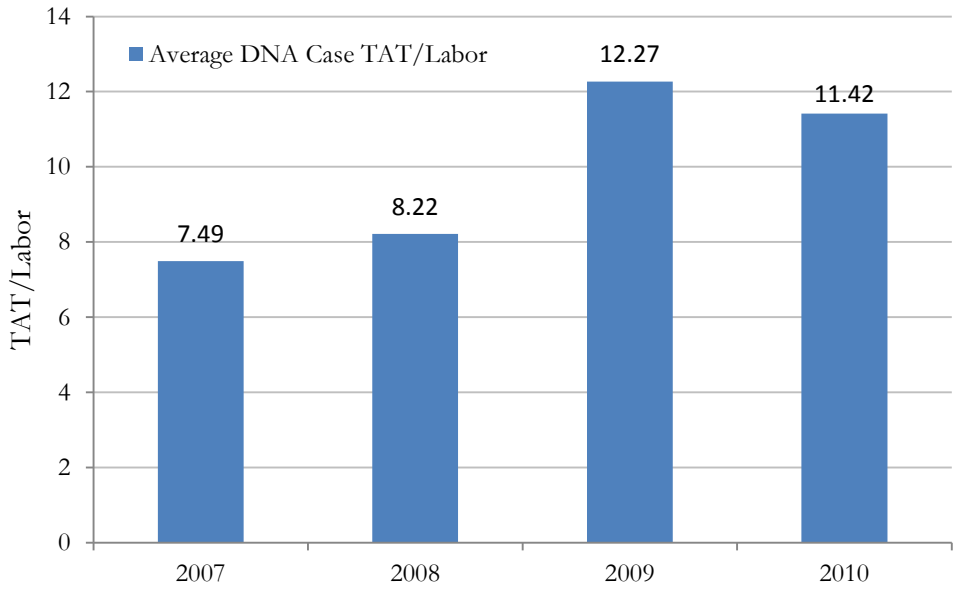


Figure 18. Efficiency Measure of DNA Turnaround Time by Budget Expenditures

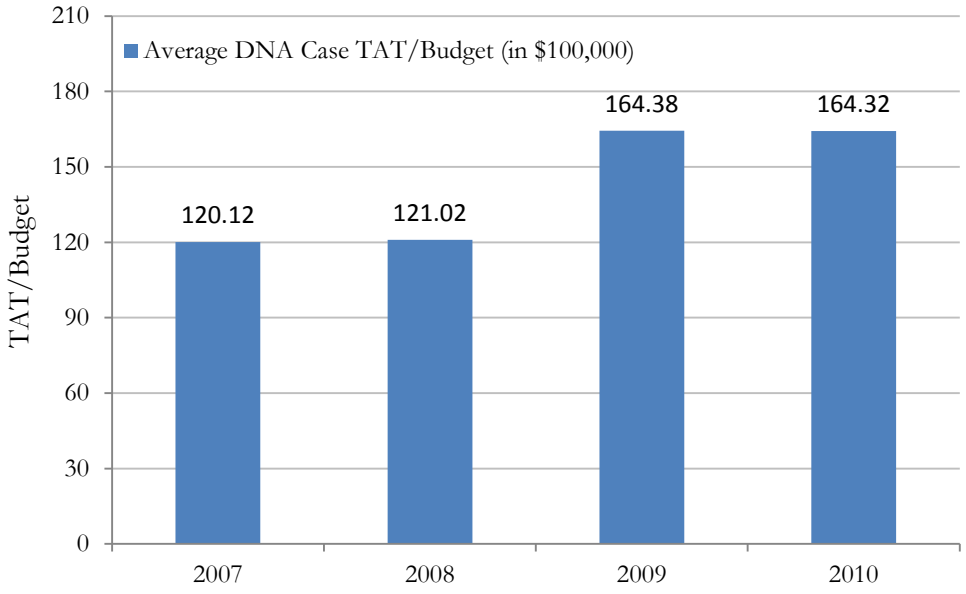
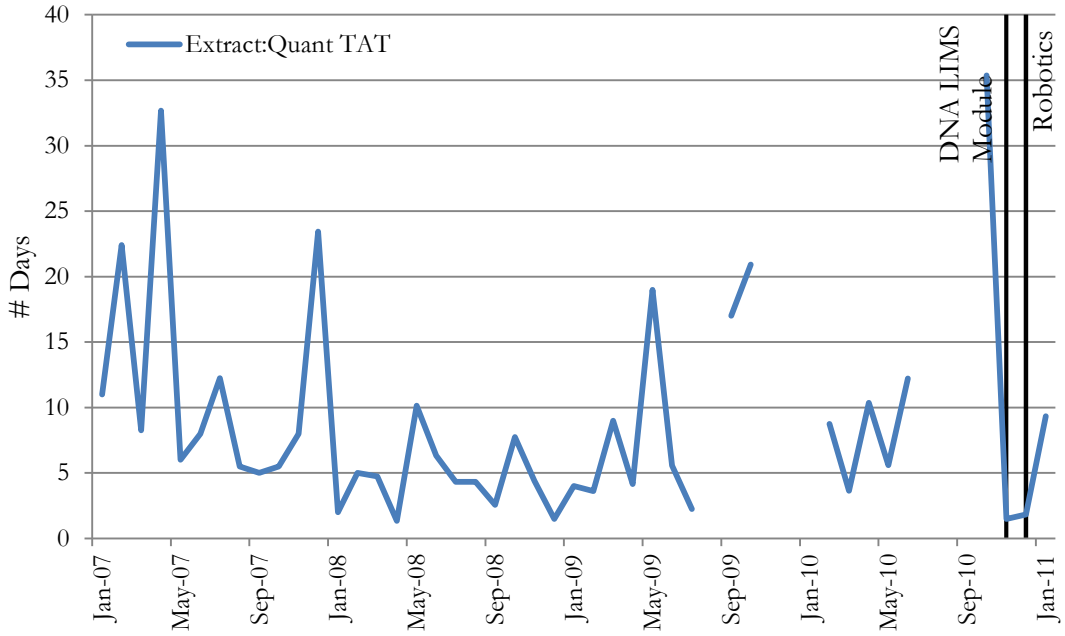
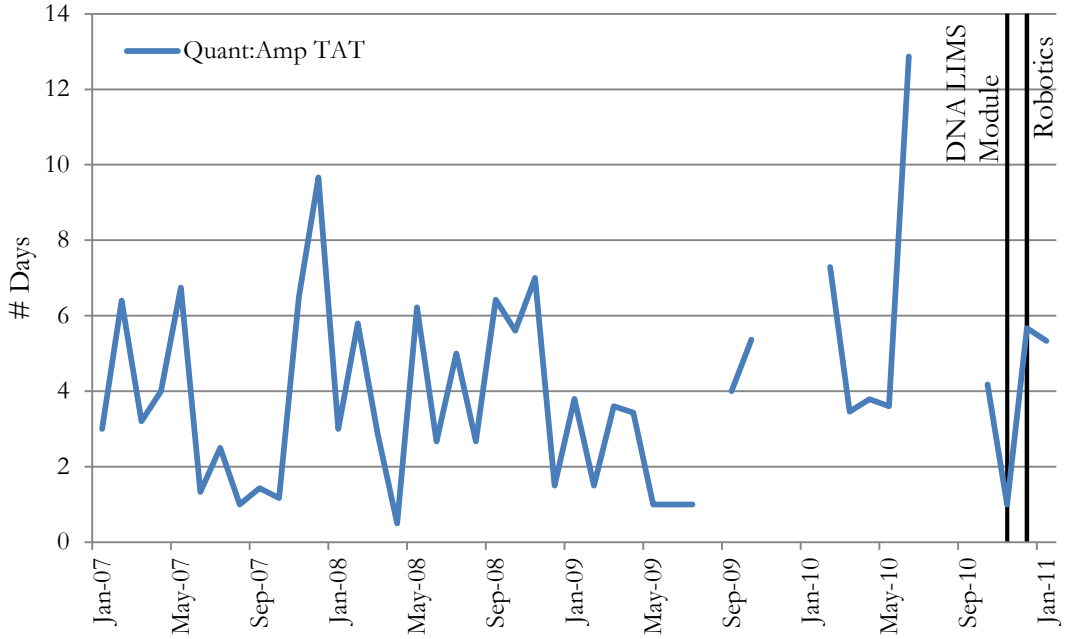


Figure 19. Stage-Level Turnaround Time: Extraction to Quantification



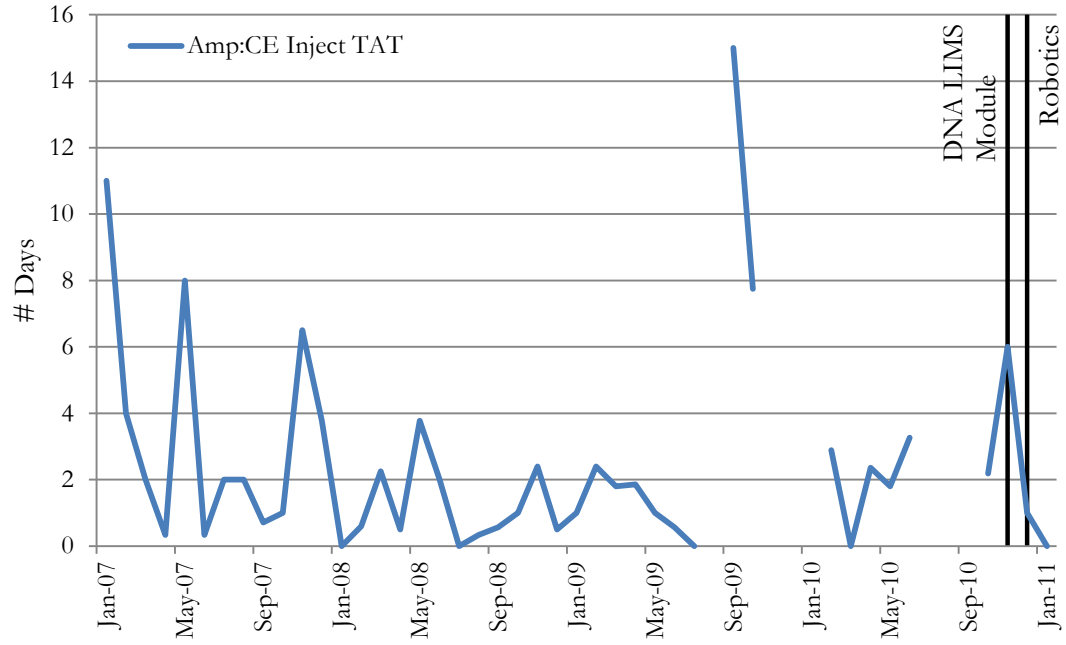
Note: Months with missing data did not have any DNA cases assigned that month.

Figure 20. Stage-Level Turnaround Time: Quantification to Amplification



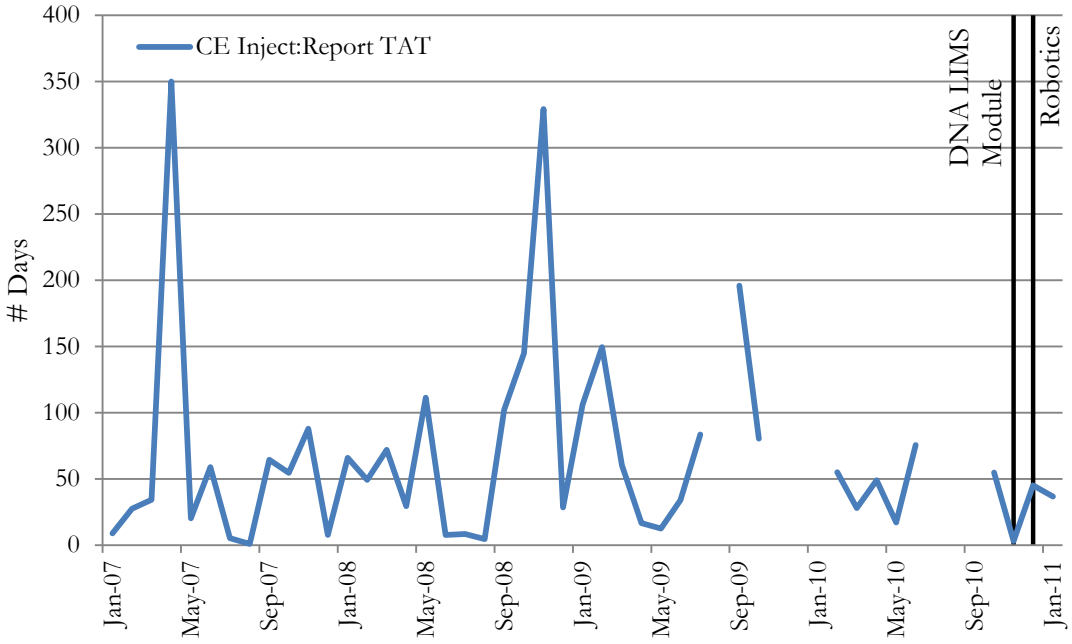
Note: Months with missing data did not have any DNA cases assigned that month.

Figure 21. Stage-Level Turnaround Time: Amplification to Capillary Electrophoresis



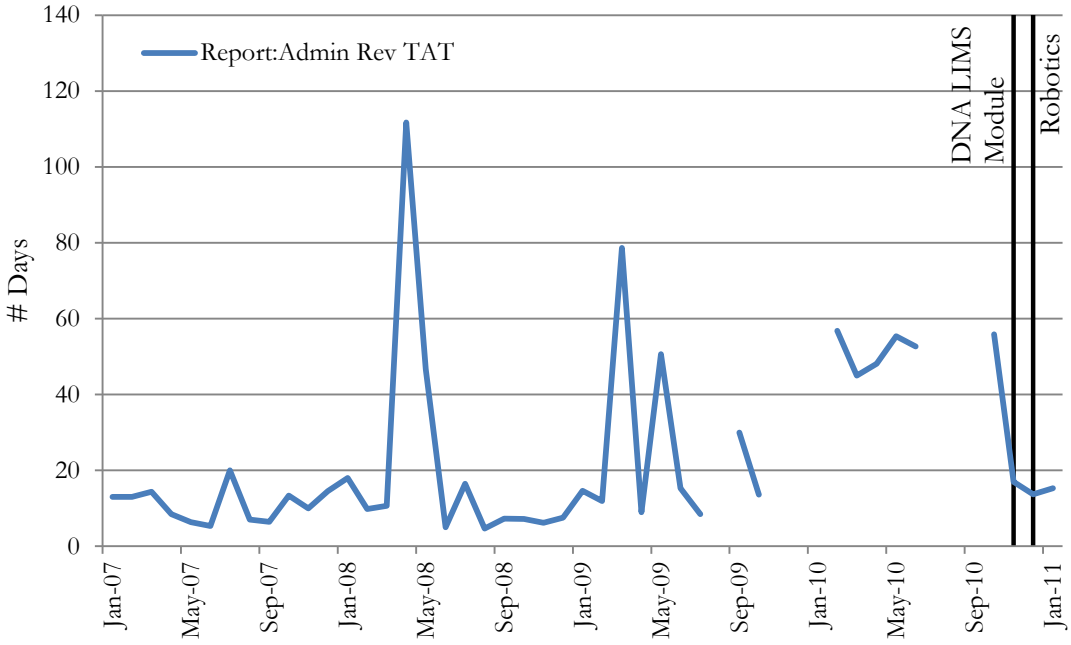
Note: Months with missing data did not have any DNA cases assigned that month.

Figure 22. Stage-Level Turnaround Time: Capillary Electrophoresis to Report



Note: Months with missing data did not have any DNA cases assigned that month.

Figure 23. Stage-Level Turnaround Time: Report to Administrative Review



Note: Months with missing data did not have any DNA cases assigned that month.

3.4.2 Pre/Post Throughput Comparison Tests

No pre/post comparison analyses were performed for Allegheny County due to the implementation delays encountered by this site. The follow-up period would only have been one month for the data collected.

3.4.3 Regression Analyses

While the implementation delays did not allow for a systematic analysis of the effects of the grant's activities on turnaround time, regression analysis can still shed light on other factors expected to influence turnaround time of casework. With this purely exploratory goal in mind, the research team conducted two negative binomial regression analyses to understand what case characteristics affect overall turnaround time for both serology and DNA casework. The researchers also attempted to analyze important factors related to stage-level turnaround time. However, there were no dates for intermediary stages of serology screening in the data. While there were intermediary stage dates for DNA analysis, the model fit diagnostics revealed problems; therefore, these regression results are not presented.

The research team included a variable for the use of the *STaCS* DNA LIMS module beginning in November 2010. However, the sample size was limited for serology cases started after this point, and there were not enough DNA cases processed after the implementation. The December 2010 implementation of robotics occurred too late to include in the model (there were not enough serology or DNA samples run after this point to compare to beforehand). Other important events in the lab were also included in the model: (1) the July 2009 move into a new laboratory facility,[37] (2) the implementation of new quantification instrumentation in June 2010 and presence of summer interns (June/July, 2010), and (3) new multichannel verification system (MVS) quality control software (October 2010) and barcode tracking protocols along with DLIMS module implementation that occurred nearly the same time (November 2010).

While there were some similarities in influential case characteristics across the serology and DNA casework, there were also variables which appeared to contribute uniquely to each (see table 4). The regression analysis did not reveal that lab events were strongly related to serology or DNA casework (however, two of the event milestones could not be examined in the DNA sample due to sample size issues explained in the note below the table). Turnaround time increased for both serology and DNA cases with more items submitted to serology, and violent offenses were associated with longer turnaround times (possibly due to more complex

[37] While all other events were coded dichotomously as 0 or 1 for before and after the implementation, cases were coded as 1 for the lab move if they began during the period the lab reported casework was affected by the lab (one month before to one month after the move for serology and one month before to four months after move for DNA).

casework). More experienced staff (criminalists for serology and DNA analysts for DNA) also had cases with longer turnaround times. This may be due to senior staff having additional supervision and management responsibilities or senior staff being assigned more complex cases. Cases with reruns or multiple submissions tended to have longer turnaround times for serology analysis, but not for DNA processing. In addition, the identification of a suspect prior to sample processing also increased turnaround time. It is unclear why this effect would be found after controlling for number of items and multiple submissions (two possible reasons a suspect might *increase* the turnaround time of a case); however, this variable may be nonetheless tapping into the effects of multiple submissions, as this site did not consistently track this in their database and it is likely that some of this information was lost. DNA cases, on the other hand, revealed a significant relationship between turnaround time and item type. Cases with sexual assault kits were more likely to have shorter turnaround times. Also, in addition to violent offense cases having longer turnaround time times, property crimes also resulted in longer time spent at the DNA level.

Table 4. Regression Results for Allegheny County

TAT Regression	**Overall Serology Case TAT**		**Overall DNA Case TAT**	
	b coefficient	*p-value*	*b coefficient*	*p-value*
Intervention: DNA LIMS Module and Confound: Barcode Tracking and MVS Quality Control	-0.22	0.08	N/A	N/A
Confound: Lab Move	-0.08	0.53	N/A	N/A
Confound: 7500 Quant Instr. + Summer Interns	-0.16	0.10	-0.19	0.29
Rerun or Multiple Submissions	2.07	<.01	0.17	0.15
Number of Serology Items	0.01	<.01	0.02	<.01
Item Type: Sexual Assault Kit	-0.02	0.73	-0.25	0.03
Item Type: Textile	0.00	0.99	-0.13	0.28
Suspect Present	0.14	0.01	-0.06	0.68
Violent Offense	0.23	<.01	0.68	0.01
Property Offense	-0.13	0.26	0.69	0.01
Criminalist/Analyst Experience	0.01	<.01	0.03	<.01

Note: Due to smaller sample size for cases with DNA work, the variables for (a) the lab move and (b) STaCS DNA LIMS module could not be tested because there was not a large enough split in cases experiencing either condition.

3.4.4 Conclusions

The Allegheny County Medical Examiner's Office Forensic Laboratory Division proposed to modify their serological screening and DNA analysis processes for sexual assault evidence. The proposed approach included a new automated sperm detection microscope, utilization of Y-STR analysis as a screening tool for male DNA in mixture samples, the use of an intelligent "genetic calculator" expert system, automatic transfer of data from the expert system to a DNA LIMS module, and the minimization of manual, repetitive liquid-handling tasks through the implementation of robotic systems.

Allegheny County was not able to implement their grant-funded interventions until the end of the study period due to a series of implementation challenges, the most inhibiting being changes to key personnel. During the study, the lab implemented three components of the new DNA process: the automated sperm-detection microscope, robotics, and DNA LIMS module. However, both the sperm-detection microscope and DNA LIMS module were removed due to technical problems, leaving only the robotics in place by the end of the evaluation. At the time of this report, the lab was still working on validating and implementing its remaining components with the exception of the Y-STR screening step, which they ultimately viewed as ill-suited to the lab's processing needs. Because Allegheny County had such delayed implementation, the research team was unable to determine whether there had been any effect of the grant program on the lab's DNA processing (due to such a modest follow-up period).

There was high variability in productivity outcome measures (i.e., throughput and turnaround time) across both cases and time, which created additional challenges in detecting patterns. Efficiency indices showed fairly similar patterns to productivity measures when graphed across time. 2010 appeared to be the most efficient year for the lab when accounting for both staff and financial resources. Further, the data show that the lab has been continuously improving its efficiency each year in regards to serology. In contrast, earlier years (2007 and 2008) were more efficient for DNA casework. However, any changes evident in the data are not due to the grant program, as implementation did not occur until the end of the study period.

Pre/post comparison tests of throughput could not be used with this site because there was an inadequate follow-up period. Regression analyses were used primarily to understand what general factors influence turnaround time (as opposed to a test of the effects of the grant program). As expected, cases with more submitted items and multiple runs or submissions tended to have longer turnaround time. In addition, violent offenses and more experienced criminalists/analysts were also associated with longer turnaround times. These findings may be due to violent offenses requiring more complex analysis and senior staff having competing management responsibilities or being assigned to more difficult cases. The

identification of a suspect also was related to a longer serology turnaround time, although it is unclear why this effect would occur after controlling for number of items and multiple submissions (two possible reasons a suspect might *increase* the turnaround time of a case). Cases with sexual assault kits were more likely to have shorter DNA processing turnaround times, possibly due to the more standardized nature of such kits. At the DNA-level, but not serology-level, property crimes often also took more time. This may be due to the nature of property crime scenes, which may be more likely to have "touch" or other low-quality and low-quantity types of DNA samples.

In conclusion, Allegheny County proposed an ambitious project involving automated procedures (through the sperm-detection microscope and robotics), expert systems, advances in data tracking with a DNA LIMS module, and a paradigm shift for the role of Y-STR analysis in sexual assault cases. Due to encountered implementation and technical challenges, the lab was unable to validate and implement the entire proposed process, but instead successfully instituted robotics into their current workflow. Unfortunately, this occurred at such a late date that the research team was unable to evaluate its effects. Although the research team's data cannot confirm such claims, the lab perceived these robotics to have a substantial effect on its ability to conduct casework, reporting that they were currently on track to produce more than twice the number of reports in 2011 compared to those produced in 2010.

4. CASE STUDY: KANSAS CITY POLICE DEPARTMENT CRIME LABORATORY

4.1 Overview of the Laboratory

The Kansas City Police Department Crime Laboratory (hereafter, Kansas City) is an accredited public crime laboratory housed within the Kansas City, Missouri, Police Department. The laboratory accepts forensic samples from Kansas City and the surrounding region.

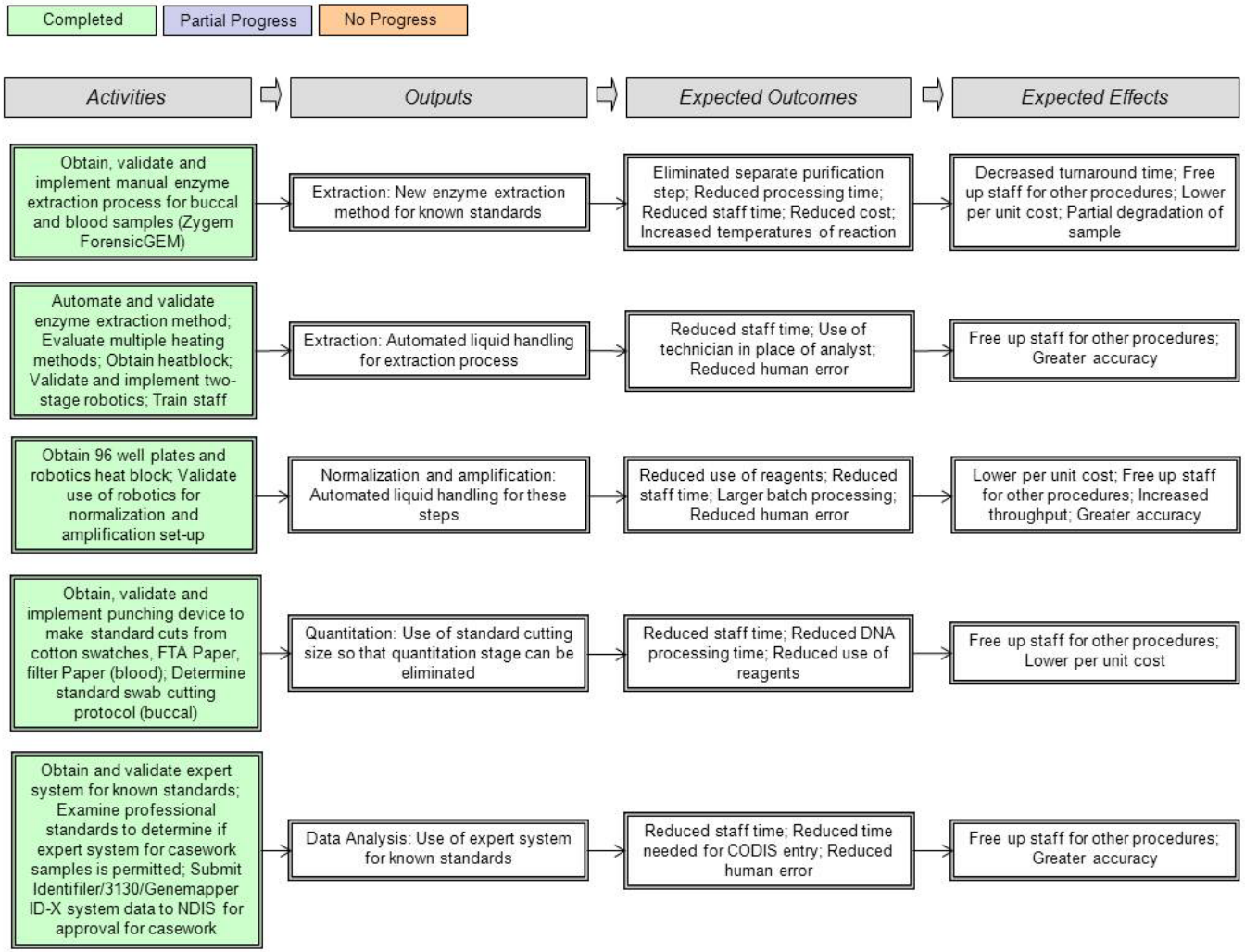
The activities proposed by Kansas City targeted the processing of known standards. Known standards are DNA samples collected directly from individuals whose identity has been confirmed (as opposed to the unknown source of "questioned" or evidence samples). These have traditionally been processed in the same manner and workflow as questioned evidence samples. However, known standards are physically more similar to samples collected for convicted offenders.[38] DNA processing data from convicted offender samples can be processed with different technologies, such as expert systems. Kansas City decided to validate the use of some of these well-accepted, convicted DNA processing technologies for use with known standards. In particular, they proposed to create a more streamlined system for processing known standards through a new sample preparation and extraction technique, automation, and expert system.

4.2 Description of Grant Goals

The NIJ awarded Kansas City a $90,000 Forensic DNA Unit Efficiency Improvement Program grant, which, in combination with Kansas City's 25 percent nonfederal match, was to fund training; purchase of robotics equipment, including a gripper arm, shaker setup, and heat block for the *Biomek 3000;* and an expert system. Kansas City described five main goals of their proposed strategy to improve the efficiency of processing "known standard" samples within their lab (see figure 24). These goals, the activities involved in achieving these goals, and the expected outcomes and impacts are described below. Kansas City expected the implementation of this new approach to result in a 29 percent increase in processing of known standards.

[38] DNA samples are sometimes collected from convicted offenders and arrestees for certain eligible crimes in order to match DNA profiles with other unknown DNA samples within CODIS. DNA profiles are typically generated from blood or buccal swabs.

Figure 24. Logic Model for Kansas City



The first goal of Kansas City's proposed strategy to increase efficiency was to implement a new enzyme extraction method, the *ZyGEM ForensicGEM*, which uses a heat-controlled enzyme, neutral proteinase from Bacillus sp.EA1, to break down proteins and release DNA. The lab obtained the new enzyme extraction method from *ZyGEM Corporation* and validated its use with a manual extraction process for both buccal and blood samples. With this new extraction technique, no downstream purification step is needed, and there is a reduced cost. The process also takes less time than the traditional extraction process (20–35 minutes compared to two days) and uses less staff time. One potential drawback of this method is that greater degradation of the sample can occur from the increased reaction temperature; however, known standards are most often high-quality samples, with a sufficient amount of DNA to obtain a full profile in spite of the increased potential for sample degradation.

Kansas City also planned to purchase accessory equipment and validate a *Biomek 3000* robot to automate the extraction process, reduce hands-on time, reduce human error, and increase accuracy in DNA analysis. Kansas City proposed to use new sample-cutting techniques to eliminate the need for quantification. A hole-punching device to cut out equally sized portions from cotton swatches would be used with blood samples, and a diagram to standardize the location of cotton swab cuts would be used for buccal swabs. The removal of quantification would mean a reduction in the use of reagents, time spent processing, and staff time, leading to lower costs and freeing up staff for other lab tasks.

Kansas City also hoped to enhance their known standards workflow by creating a workflow separate from other casework samples and automating the amplification setup process using a *Biomek 3000* robot with 96-well plates. Using larger well plates permits larger batch processing, and Kansas City expected this to reduce their use of reagents and reduce human error as well as increase the number of samples amplified at one time. This it was thought would lead to lower cost, greater accuracy, and greater throughput.

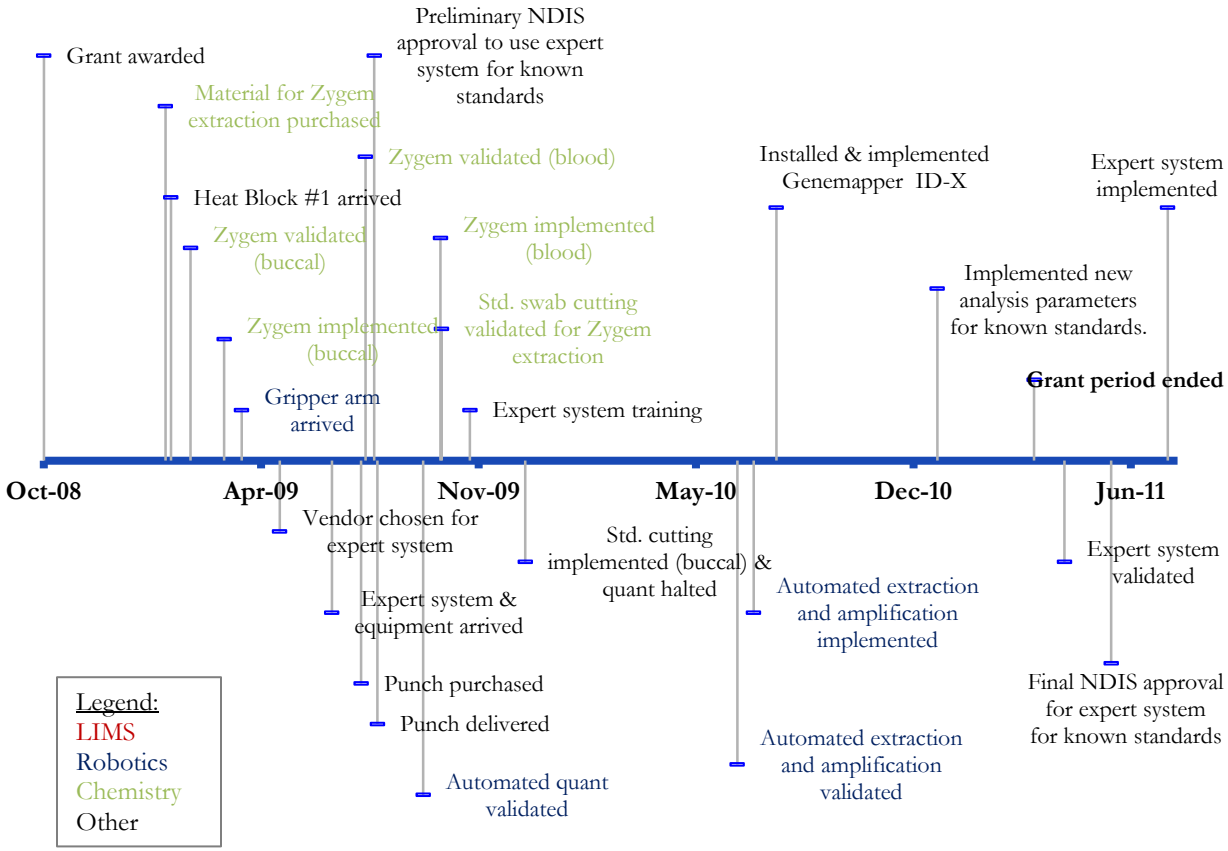
An expert system was also included as part of Kansas City's approach to improving efficiency. Because expert systems are approved by the FBI for use with convicted offender and arrestee samples, Kansas City reasoned that expert systems should also be permissible for known standard samples which are collected in similar ways. However, since this was a new approach they would need to obtain approval from the FBI National DNA Index System (NDIS) board for use with known samples.[39] With an expert system, only one analyst would be needed for conducting a technical review of the DNA data rather than having two analysts independently review the data. Using an expert system was expected to reduce staff time spent on data review and decrease human error.

[39] NDIS board approval for use of expert systems with data from known standards would allow other public laboratories interested in using expert system technology to do so while maintaining compliance with NDIS Board policies.

4.3 Implementation Findings

4.3.1 Implementation Description

Figure 25. Kansas City Implementation Timeline



Kansas City concluded their grant period on March 31, 2011, 29 months after the beginning of the grant period. Kansas City successfully completed all goals of their grant proposal. Implementation milestones are shown in Figure 25. The lab validated and implemented the new enzyme extraction method, sample cutting procedures, expert system, and to automate extraction, amplification, and data review.

The new extraction process was validated with a *Biomek 3000* robot, and *Genemapper ID-X* was purchased for use as an expert system. The lab validated the expert system with over 1,200 known standard samples according to the NDIS guidelines, trained staff on its

use, and worked with the NDIS board to obtain approval for the use of expert systems with known standards. With the greater consistency in sample cutting, Kansas City was able to discontinue the quantification step, as they no longer needed to measure the amount of DNA in their sample. The lab did not receive approval from the NDIS Board for use of the expert system for actual casework known samples until after the grant period had ended. However, once approval was obtained, the lab instituted the use of the expert system with casework known samples and has been using the *fully* implemented process since that time.

4.3.2 Implementation Challenges

Kansas City did not encounter many implementation challenges. The project manager of Kansas City's project said the largest implementation delay was in receiving the correct heat block equipment for the *Biomek 3000* and obtaining approval from NDIS for use of the expert system. The original heat block Kansas City ordered did not have the required two heating settings and therefore needed to be replaced with a more suitable heat block. The lab encountered delays in obtaining approval from NIJ for a new heat block for the robotics equipment and the subsequent construction of this heat block by the vendor. NDIS approval was not received until June 2011, three months after the close of the grant period. However, by August, the lab was using the expert system for known standards.

Kansas City needed to adjust their laboratory workflow and reorganize staff in order to allow known standards to be batch-analyzed together, as opposed to being batch-analyzed with the rest of the case evidence samples. One forensic technician was able to process all of the known standard samples, while another analyst conducted the second technical review of the data (since the expert system performed the first review). Other reported challenges were the additional time needed for developing procedure manuals and integrating these into existing lab documents and their LIMS. In addition, it was difficult to find time to work on validation studies with competing obligations to perform casework and work on a separate LIMS development project.

Like in Allegheny County, there were also some personnel shifts which occurred. The grant project manager was on personal leave for a substantial portion of 2010. However, this did not have a large impact on the project, as the majority of the activities were completed by that time and knowledge had been successfully shared with other staff.

4.3.3 Final Perceptions

Kansas City saw many benefits of the NIJ grant program. They perceived the most important impacts to be increased throughput, decreased turnaround time, and the establishment of two separate workflows: one for "questioned" forensic evidence samples and one for known standards. Lessons learned from this new workflow also influenced traditional processing. For example, the lab reported they were shifting toward using technicians more and trying to

use an assembly-line framework for other types of casework evidence. Kansas City also felt that other labs would be able to learn from their experience and institute similar procedures to increase efficient processing of known standards.

The interviewed point of contact felt that the biggest challenge of the project was to find the time and resources to conduct validation studies while still performing normal casework responsibilities. Incorporating the new process into the existing lab's workflow and juggling other lab changes was another identified difficulty. For example, one particular challenge involved changing the LIMS infrastructure while implementing new grant interventions. Kansas City also reported some confusion over a few NIJ requirements and thought that more clear expectations from the outset would help future grantees. Overall, the lab felt the grant project was beneficial, and the lab plans to continue using the newly developed process.

4.4 Outcome Findings

The following section describes the data used to assess the outcomes of the NIJ Forensic DNA Unit Efficiency Improvement Program on Kansas City and changes in productivity and efficiency at Kansas City.

4.4.1 Descriptive Statistics and Trend Analysis

Kansas City's analytic file consisted of 3,173 known standard DNA samples related to 1,057 cases. Known standards comprised nearly one-third (31.1 percent) of the overall DNA casework samples analyzed at the lab during the study period. While the majority of known standards were buccal swab samples, a sizable number (30.9 percent) were blood samples. The proportion of blood samples did not substantially change over time.

Over two-thirds of known standard samples were related to violent offenses (68.9 percent), including about one-quarter (26.4 percent) that were for homicide cases and a little over one-third (36.7 percent) for sexual assault cases. A smaller proportion of known standard samples were for property crimes (12.5 percent), drug crimes (6.6 percent), and other types of crime (13.8 percent).

A suspect was identified for the majority (69.0 percent) of samples, and more than half (58.4 percent) of samples were rated at the highest priority level (followed by 37.4 at the second-highest priority level and 4.2 percent listed as lower priority). Nearly half (49.2 percent) of these samples were related to cases that had multiple submissions,[40] and a small

[40] This would impact the dates of stages, which are reported only for the case in its entirety—specifically, technical review and report dates for samples analyzed after the new LIMS began tracking these by case instead of sample.

number (4.3 percent)[41] involved reruns of the samples at one of the stages of processing. Eleven individuals were listed as analysts responsible for sample processing.

Across the four years, the median throughput was 57 samples per month (see table 5 for productivity and efficiency estimates). The median turnaround time for the entire processing of a sample (from assignment to report) was 57 days. The period between submission and assignment was around 75 days, indicating that samples have long wait periods before being assigned. Processing stages varied in turnaround times between one and 10 days, with the stage between amplification and injection taking the least amount of time and the stage between assignment and extraction taking the longest.

Statistics for efficiency indices (throughput and turnaround time divided by annual labor counts and budget expenditure estimates [in $100,000 units]) are also shown in table 5. When accounting for staff resources, the lab completed about 7.30 samples per month per analyst during the four-year period.

[41] While reruns at every stage could be coded for all cases assigned after March 2010, only reinjections were able to be coded for cases before that.

Table 5. Kansas City Throughput and Turnaround Time Outcomes

Known Standards (N = 3,173)		Productivity/Labor	Productivity/Budget	Cleaned Productivity	Raw Productivity
Overall Outcomes					
Sample Turnaround Time	Mean	10.46	36.77	82.34	91.02
Assignment–Report	Median	7.27	23.04	57.00	56.00
	Std. Dev.	8.90	35.35	70.34	136.76
	Range	(.46, 65.93)	(1.31, 247.94)	(4, 493)	(-308, 7334)
Sample Throughput	Mean	7.66	25.36	N/A	60.39
	Median	7.30	18.59	N/A	57.00
	Std. Dev.	4.69	24.16	N/A	38.21
	Range	(0.52, 30.92)	(1.86, 148.27)	N/A	(4, 250)
Stage-Level Turnaround Time					
Submission–Assignment	Mean	17.72	50.06	134.96	149.70
	Median	9.00	22.86	75.00	89.00
	Std. Dev.	39.46	118.91	287.72	284.81
	Range	(0, 477.11)	(0, 1915.88)	(0, 3300)	(-1, 3300)
Assignment–Extraction	Mean	3.61	12.44	28.41	39.47
	Median	1.31	3.86	10.62	10.00
	Std. Dev.	6.29	21.35	49.33	104.75
	Range	(0, 50.02)	(0, 197.85)	(0, 350)	(-53, 1321)
Extraction–Amplification	Mean	0.75	2.24	5.82	1.94
	Median	0.52	1.31	4.00	5.00
	Std. Dev.	0.93	3.51	7.16	428.51
	Range	(0, 9.39)	(0, 46.75)	(0, 72)	(-39439, 687)

Stage-Level Turnaround Time		Productivity/Labor	Productivity/Budget	Cleaned Productivity	Raw Productivity
Amplification–CE Injection	Mean	0.27	1.03	2.16	3.74
	Median	0.12	0.33	1.00	1.00
	Std. Dev.	0.58	2.67	4.64	431.72
	Range	(0, 5.07)	(0, 24.50)	(0, 41)	(-2191, 39445)
CE Injection–Interpretation	Mean	1.59	5.13	12.43	40.03
	Median	0.58	1.57	5.00	6.00
	Std. Dev.	3.24	10.48	24.90	141.13
	Range	(0, 42.07)	(0, 111.86)	(0, 304)	(-1091, 2207)
Interpretation–Tech Review	Mean	2.71	11.31	21.69	-12.91
	Median	0.52	1.57	4.00	1.34
	Std. Dev.	5.18	23.15	41.78	131.09
	Range	(0, 44.96)	(0, 217.17)	(0, 363)	(-2550, 1097)
Tech Review–Report	Mean	1.63	5.18	12.53	14.29
	Median	1.04	2.94	8.00	8.00
	Std. Dev.	1.86	6.57	14.08	87.59
	Range	(0, 23.75)	(0, 59.46)	(0, 192)	(-345, 7312)

Notes: Labor is defined as the number of staff reported for that year. Budget is defined as the annual DNA unit budget in $100,000 units. Turnaround time is reported in number of days.

Four years of data reveal substantial variability in both throughput and turnaround time across the evaluation period (see figures 27 and 30). In addition, there are some spikes present in the data, including a large increase in samples completed in October 2010 and a longer average turnaround time for the month of March 2010. The increase in turnaround time in March 2010 and subsequent decrease in throughput in April 2010 are likely due to the switch to a new LIMS, during which time the lab temporarily halted DNA work to facilitate the transition. The research team and the laboratory are unaware of any event that could explain the large spike in throughput in October 2010.

There are four main intervention points shown on the graphs, including the implementation into casework of (1) *ZyGEM* buccal extraction, (2) *ZyGEM* blood extraction and automated quantification with robotics, (3) standard cutting practices, and (4) automated extraction and amplification with robotics. The expert system could not be examined, because its implementation occurred after the data collection. Figure 27 illustrates an increase in completed samples after robotics were implemented in July 2010 (even disregarding the large spike in October 2010). However, there is no clear visible pattern in relation to the other, earlier implementation milestones. There appears to be an increase in samples assigned[42] in 2010 compared to earlier years, although it is unclear whether this is related to any sort of event in the lab or to random variability in sample submissions (see figure 26).

Figures 28–29 show the annual *efficiency* indices for throughput.[43] Overall, 2010 appeared to be the most "efficient" year in terms of what was produced with available resources. Although DNA casework staff fluctuated minimally across the years, the greater number of cases completed in 2010 helped to improve the efficiency ratio. This 2010 productivity improvement was also obtained with lower budget expenditures, further improving the efficiency ratio when accounting for financial resources. However, this pattern may be deceiving, since the site reported that some supplies and equipment for 2010 were preordered in 2009.

Monthly measurements of overall sample turnaround time do not reveal a clear improvement in time spent processing known standards after grant interventions are implemented (figure 30). *Efficiency* indices of turnaround time do not show a strong pattern when turnaround time is divided by labor counts (figure 31). However, 2010 again appears to be the most efficient year when budget expenditures are taken into account, due primarily to a substantially lower budget in 2010 (figure 32).

[42] Completed cases are not necessarily the same cases as those started each month. Started cases are matched to the month in which a case was assigned, while completed cases are assigned to the month in which the case was completed.

[43] Estimates are provided by year because resource indicators were assessed on an annual basis.

Turnaround time measures have different patterns depending on the particular stage of processing (see figures 33–48). Across all stages, the large variability in turnaround time makes it difficult to detect effects of the implemented interventions. The only graphs that visibly show reductions in turnaround time potentially due to the grant are (1) figure 33, which shows a decline in time between assignment and extraction completion[44] for nearly a year after the implementation of *ZyGEM* extraction for buccal swabs, and (2) figure 36, which shows reduced time between extraction completion and amplification completion after automated quantification is implemented and continuing through the use of standard cutting procedures and automated amplification. Figure 46 also shows a decrease in turnaround time between technical review[45] and report completion; however, this is likely not due to the grant since no interventions targeted the report-writing stage. The same spike in spring 2010 appears in all stages of analysis except for the time between extraction and amplification; again, this is likely due to the LIMS transition and related halting of casework.

There was an increase in turnaround time between data interpretation completion and the technical review (figure 45). However, this is likely an artifact of the data structure. In March 2010, the lab changed LIMS databases. The new LIMS only tracks technical review dates at the *case* level, while the original data system tracked at the *sample* level. Because cases can have multiple samples and multiple submissions, a case's technical review could occur much later than the technical review for an individual sample. Therefore, the increase seen in figure 45 appears to be a data anomaly and is not indicative of true change across the period. Therefore, efficiency indices are not presented for this stage.

Efficiency findings varied by the stage of processing. In terms of labor resources, 2008 was a less efficient year for the first four stages. The year of 2009 was the least efficient year for the stages between capillary electrophoresis and report (with the exception of the interpretation to technical review stage, which cannot be validly interpreted due to data issues). Interestingly, budgetary efficiency indices were often more stable than the labor efficiency indices, and sometimes conflicted with the results of the labor efficiency indices. Since the productivity measures are the same for each set of efficiency indices, these conflicting results are due to differences in labor and financial resources (i.e., a year with more financial resources may not necessarily have more staff resources). Overall, 2009 was often the least efficient year when taking into account budget expenditures, with the exception of the first two stages.

[44] It is important to note that this is an imperfect measure of extraction turnaround time because it is unclear what proportion of the time given is wait time after a sample is assigned but before work is done.

[45] Due to inconsistencies in reporting practices, the date listed for technical review may be the date started or date completed.

Figure 26. Monthly Number of Samples Started

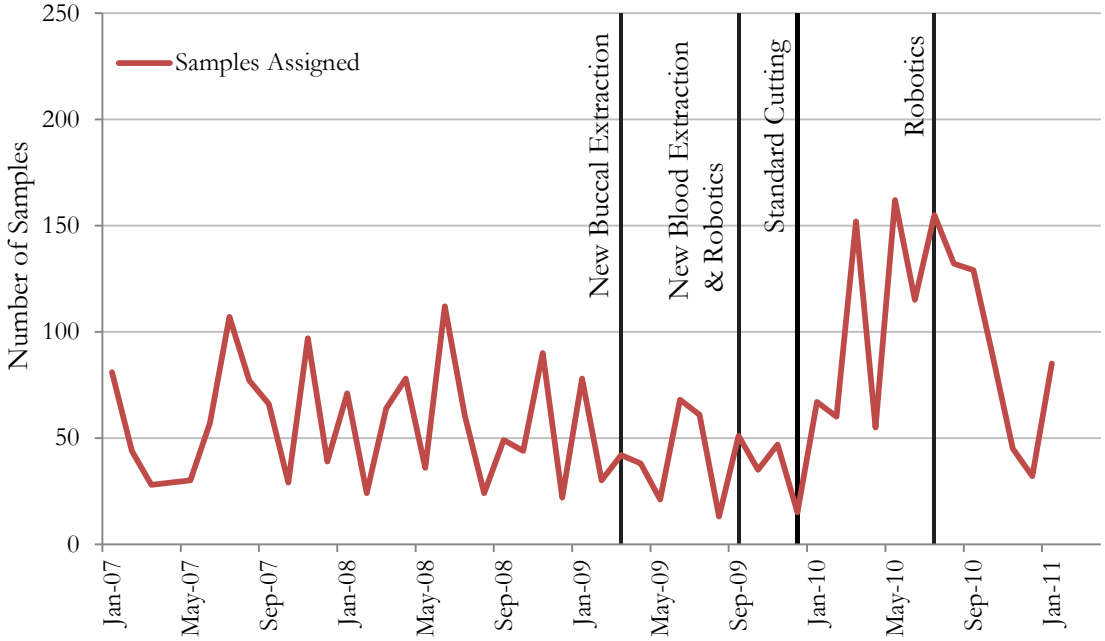


Figure 27. Monthly Throughput of Samples

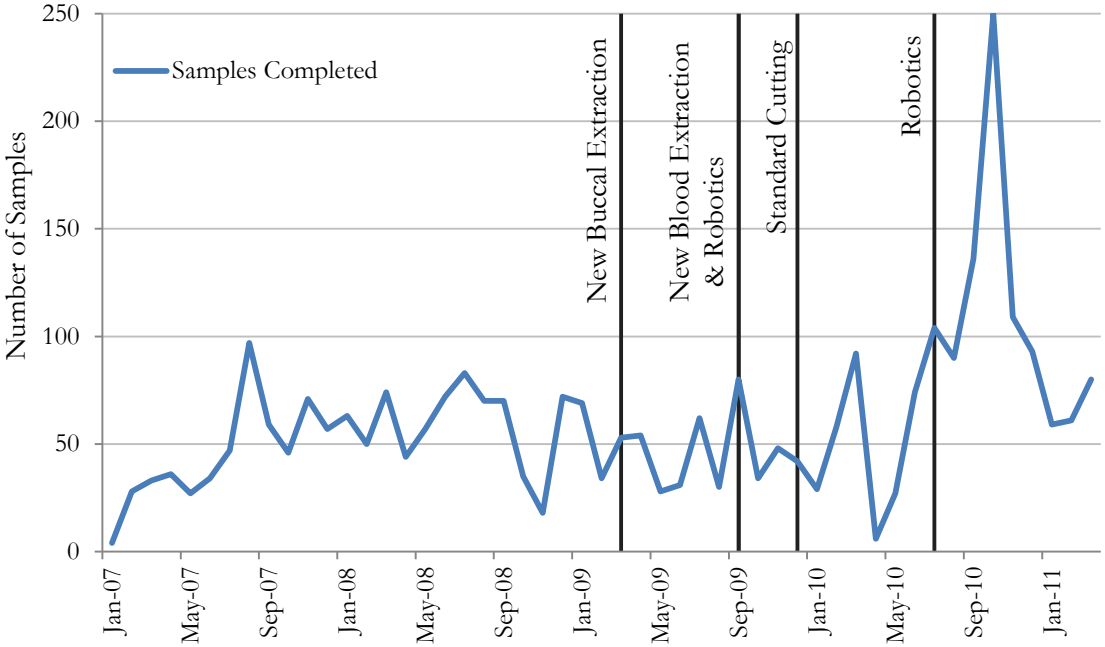


Figure 28. Efficiency Measure of Monthly Throughput by Labor

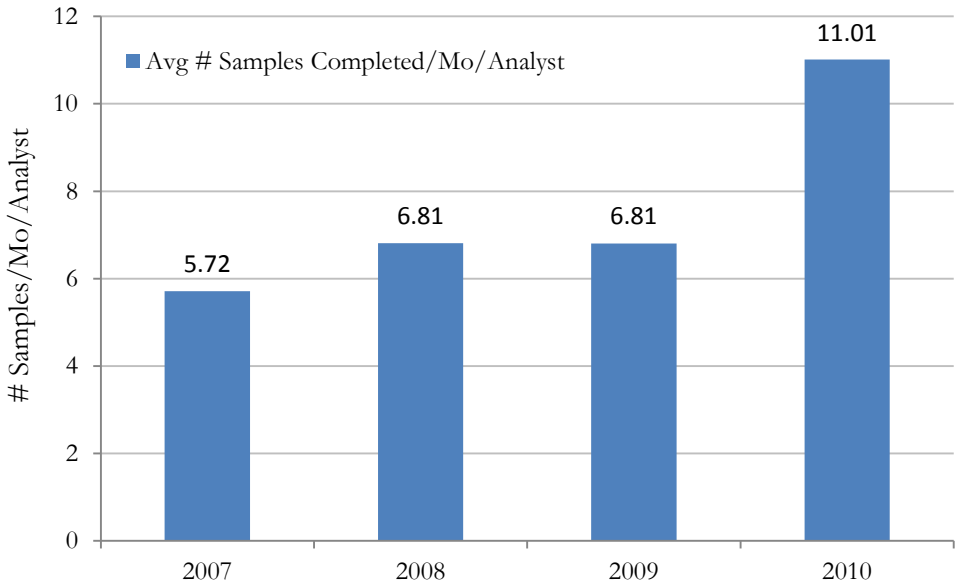
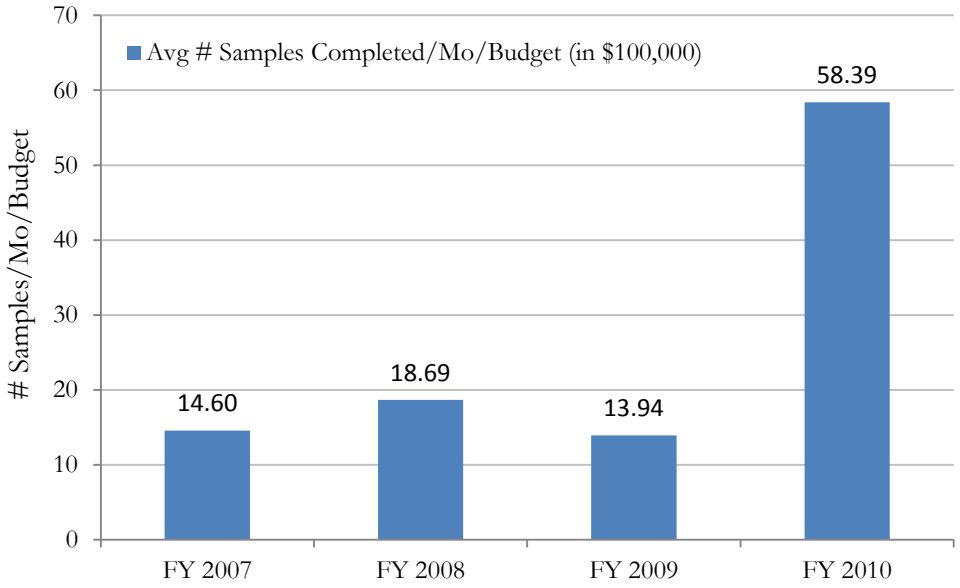


Figure 29. Efficiency Measure of Monthly Throughput by Budget Expenditures



Note: Because case processing data were not available for the last three months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but cases are only provided for 75 percent of the year.

Figure 30. Sample Turnaround Time

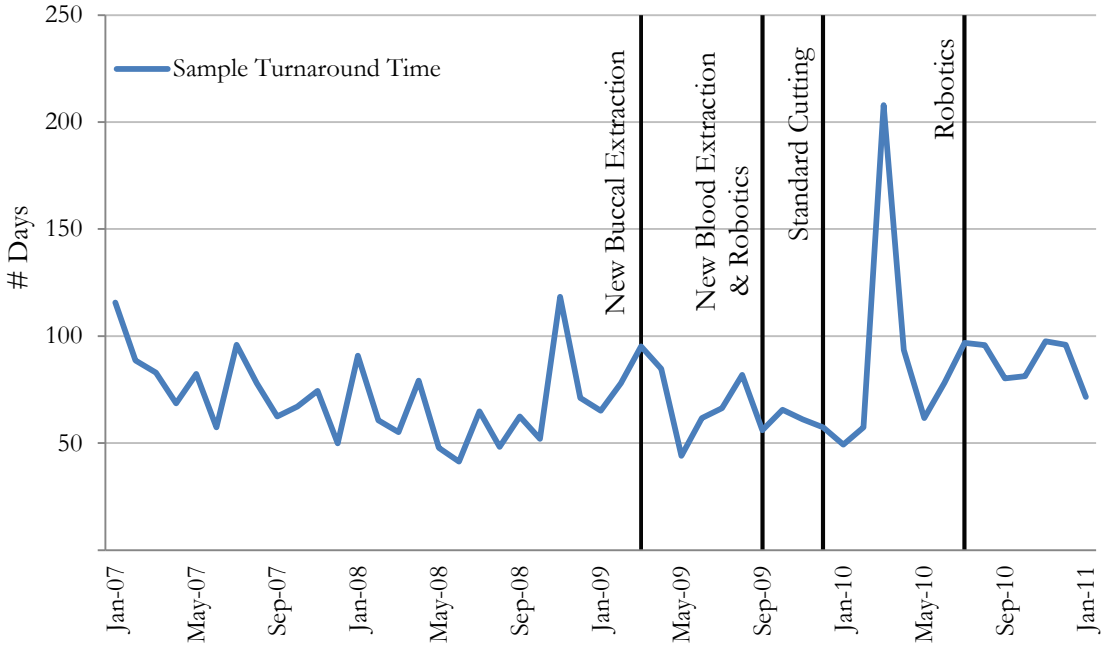


Figure 31. Efficiency Measure of Turnaround Time by Labor

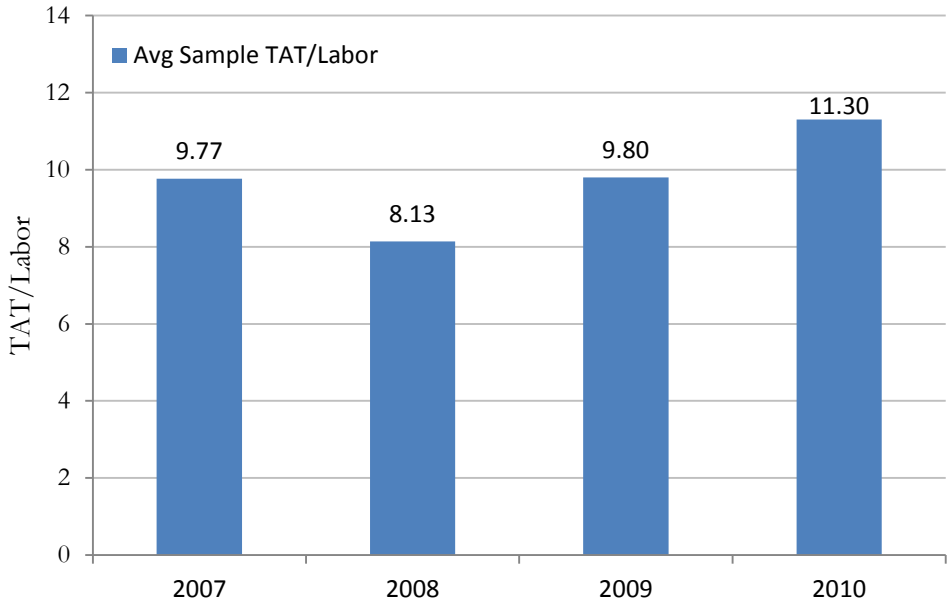
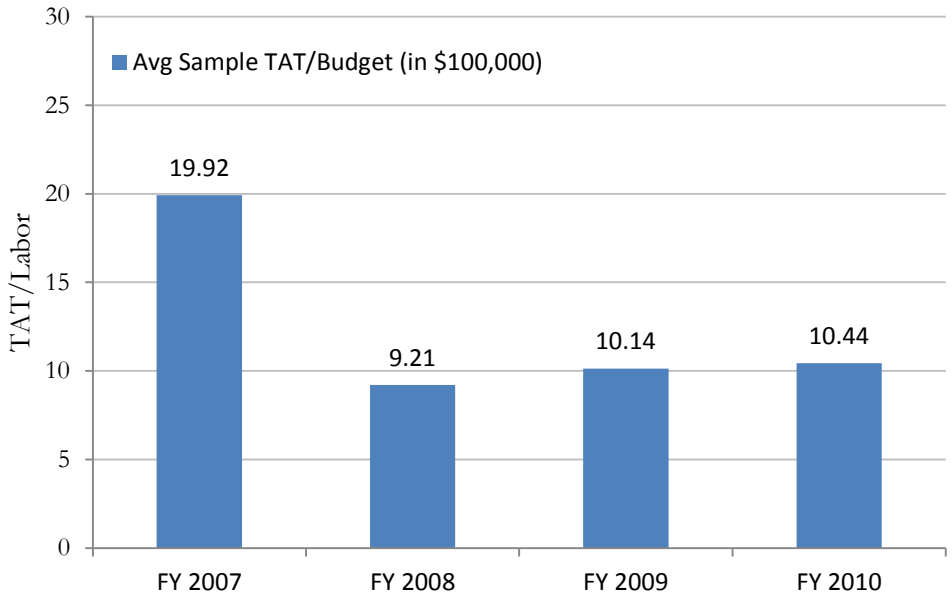


Figure 32. Efficiency Measure of Turnaround Time by Budget Expenditures



Note: Because case processing data were not available for the last three months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 75 percent of the year.

Figure 33. Stage-Level Turnaround Time: Assignment to Extraction

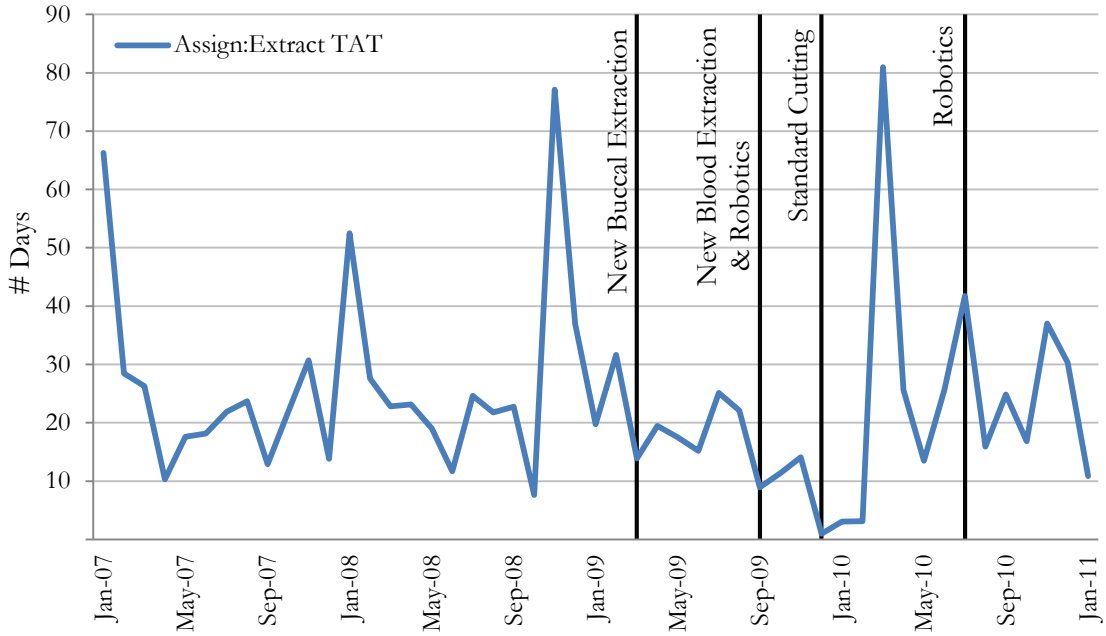


Figure 34. Stage-Level Efficiency Measure by Labor: Assignment to Extraction

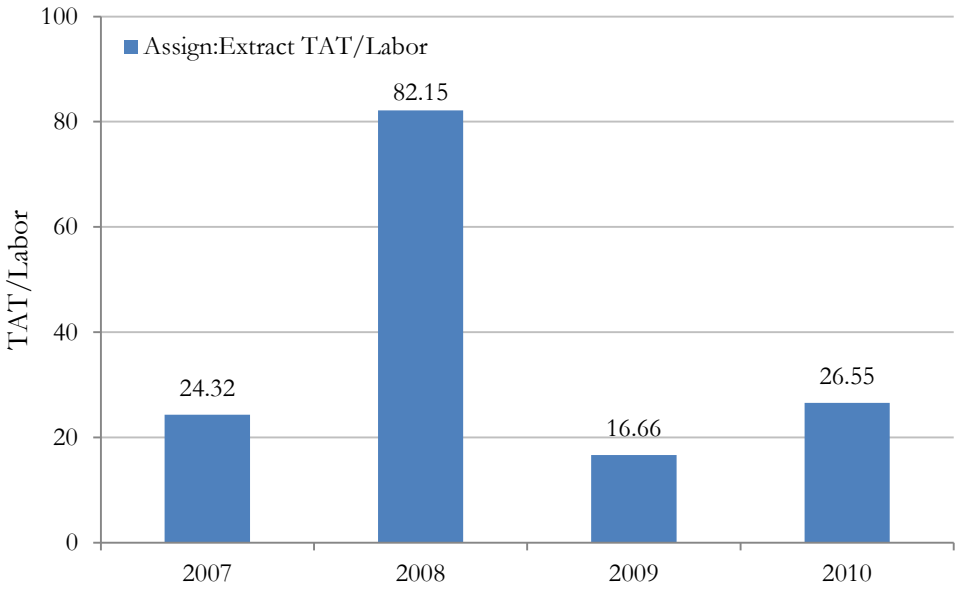
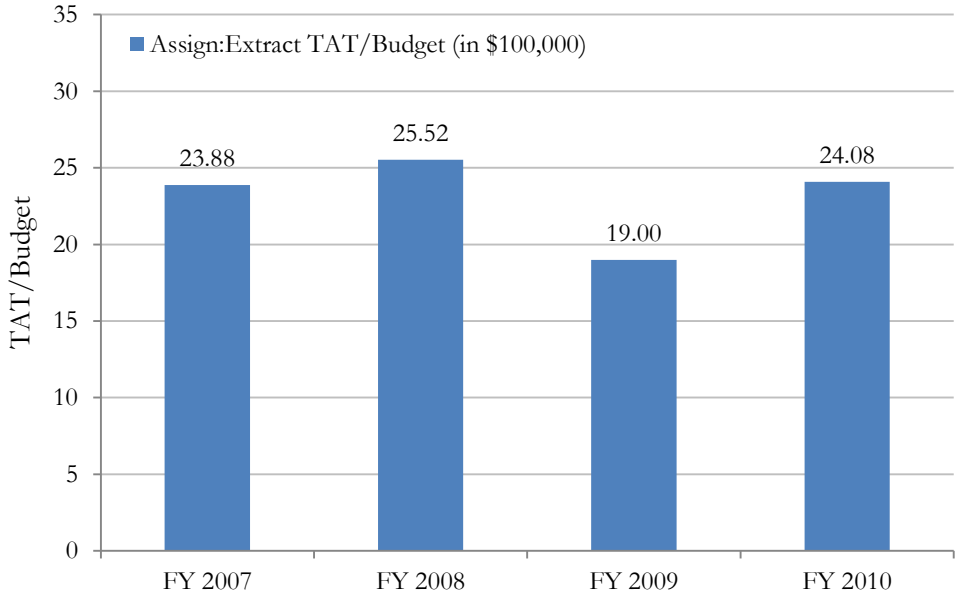


Figure 35. Stage-Level Efficiency Measure by Budget: Assignment to Extraction



Note: Because case processing data were not available for the last three months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 75 percent of the year.

Figure 36. Stage-Level Turnaround Time: Extraction to Amplification

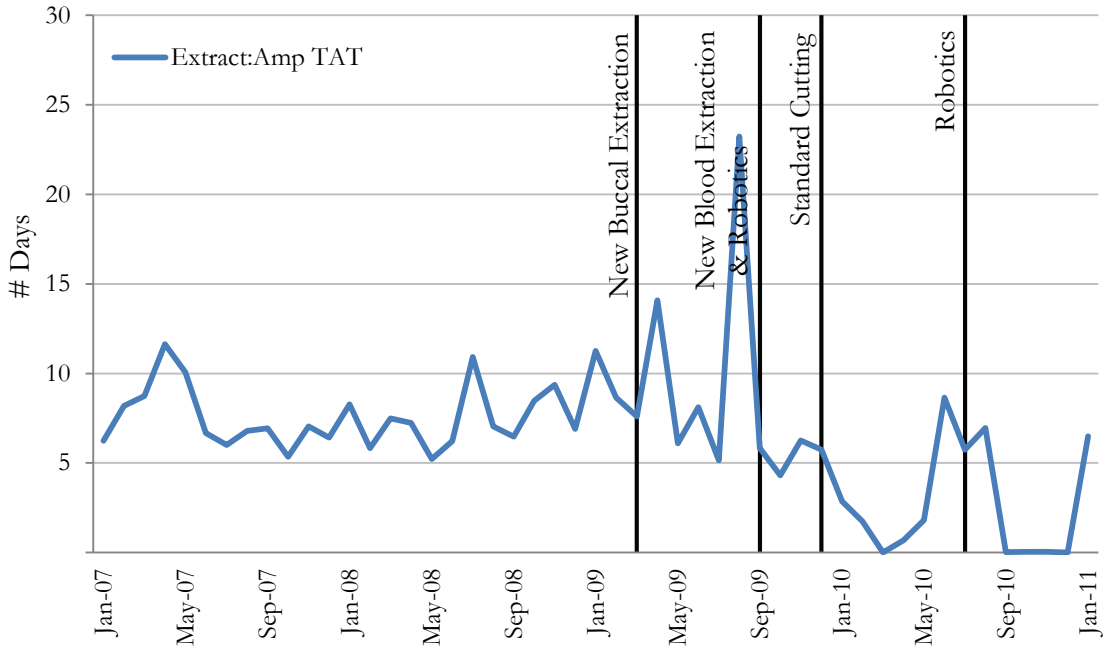


Figure 37. Stage-Level Efficiency by Labor: Extraction to Amplification

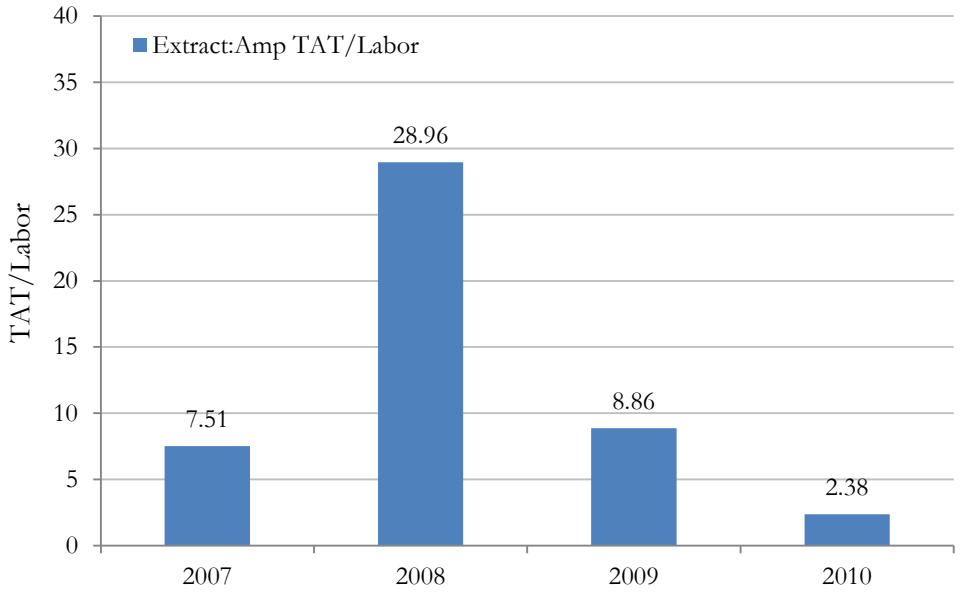
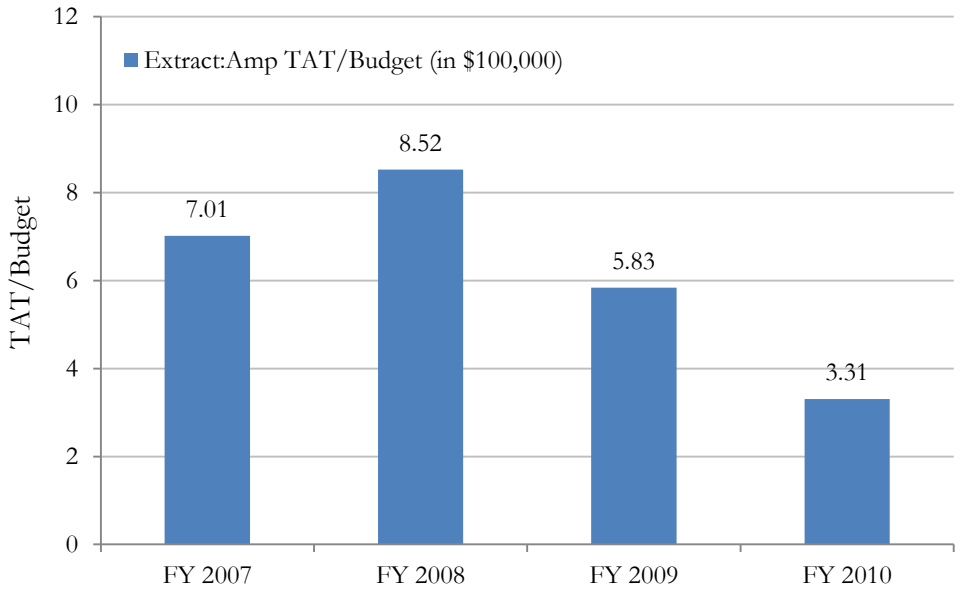


Figure 38. Stage-Level Efficiency Measure by Budget: Extraction to Amplification



Note: Because case processing data were not available for the last three months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 75 percent of the year.

Figure 39. Stage-Level Turnaround Time: Amplification to CE Injection

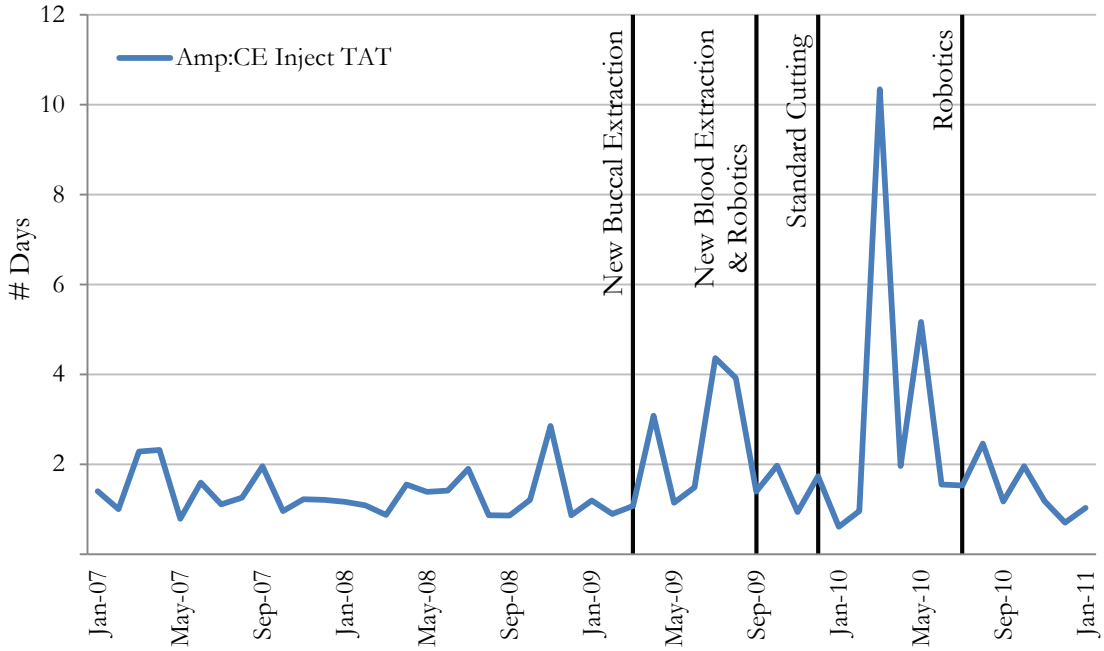


Figure 40. Stage-Level Efficiency by Labor: Amplification to Capillary Electrophoresis

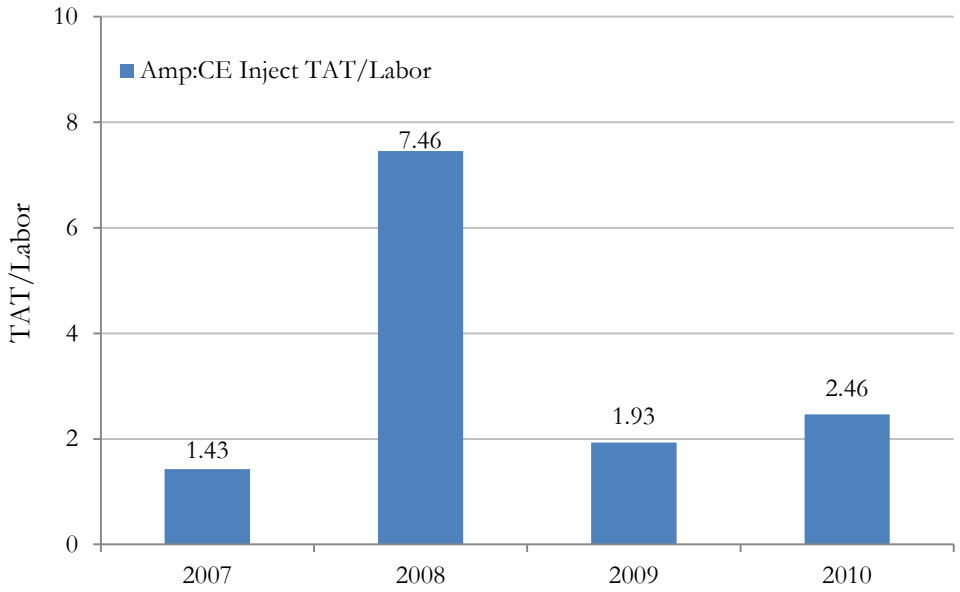
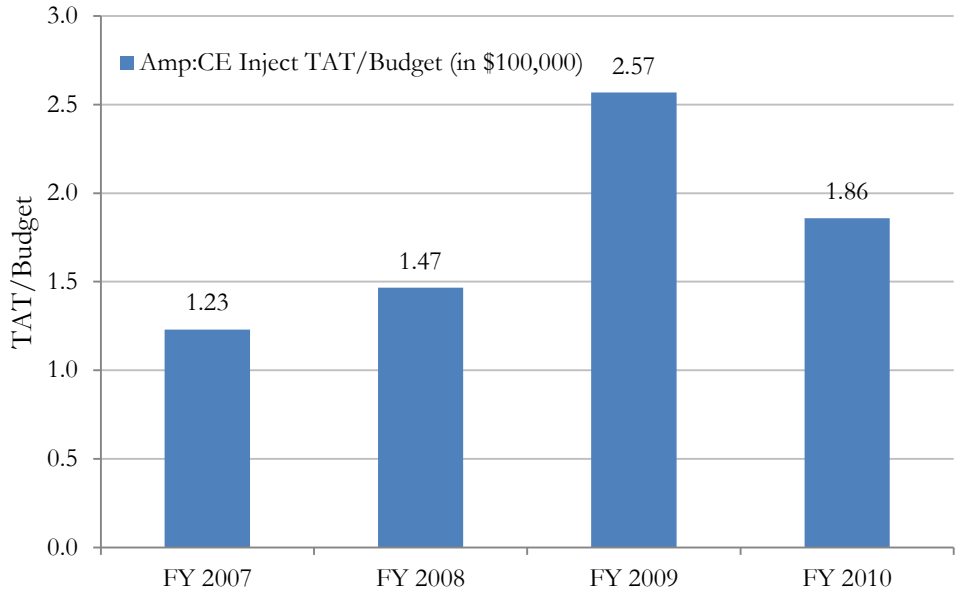


Figure 41. Stage-Level Efficiency Measure by Budget: Amplification to Capillary Electrophoresis



Note: Because case processing data were not available for the last three months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 75 percent of the year.

Figure 42. Stage-Level Turnaround Time: CE Injection to Interpretation

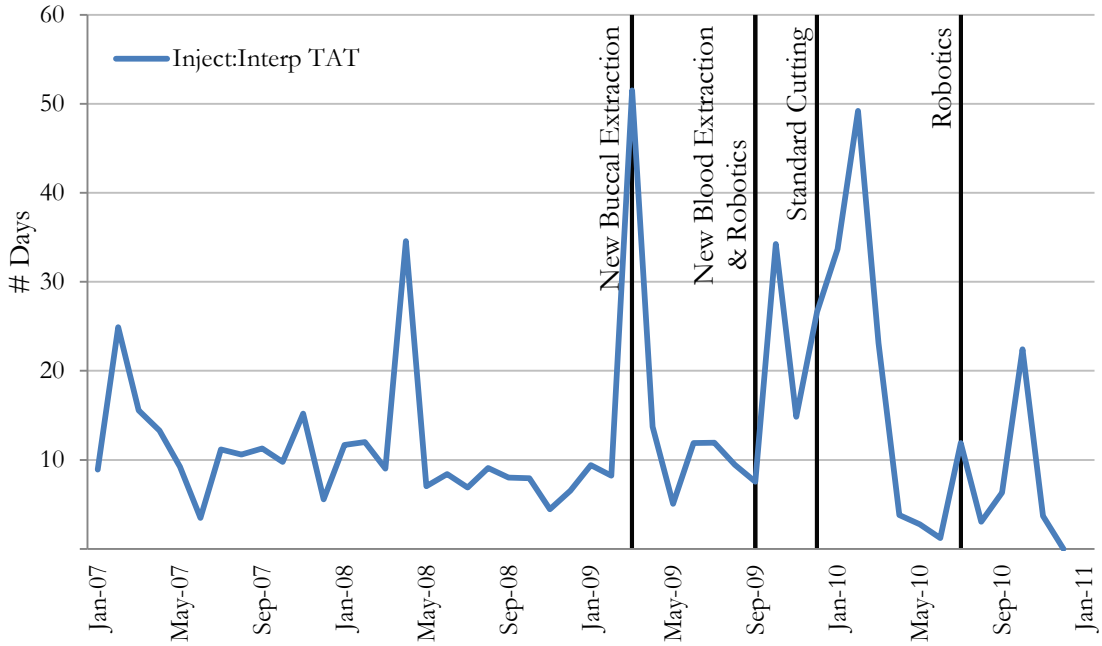


Figure 43. Stage-Level Efficiency Measure by Labor: Capillary Electrophoresis to Interpretation

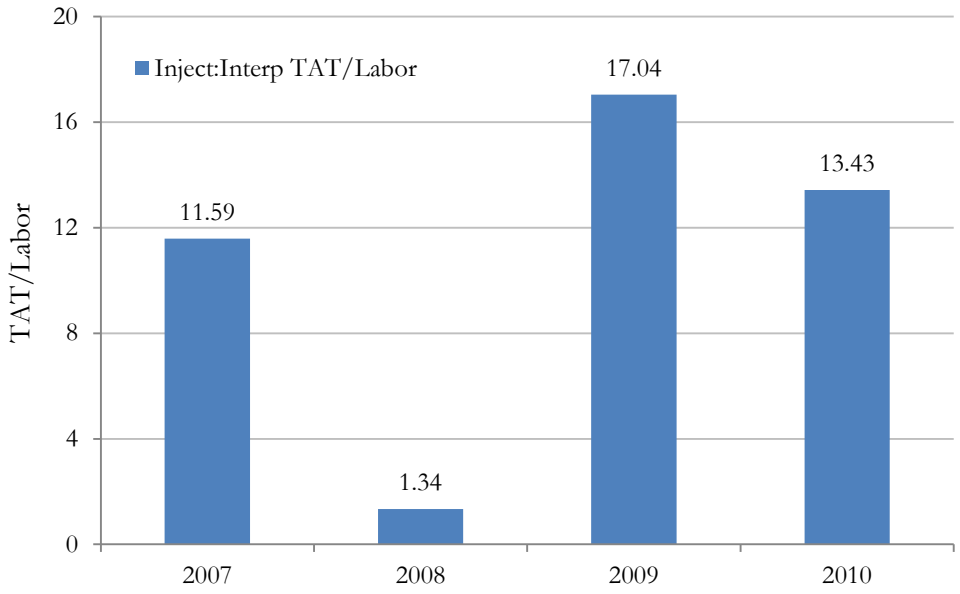
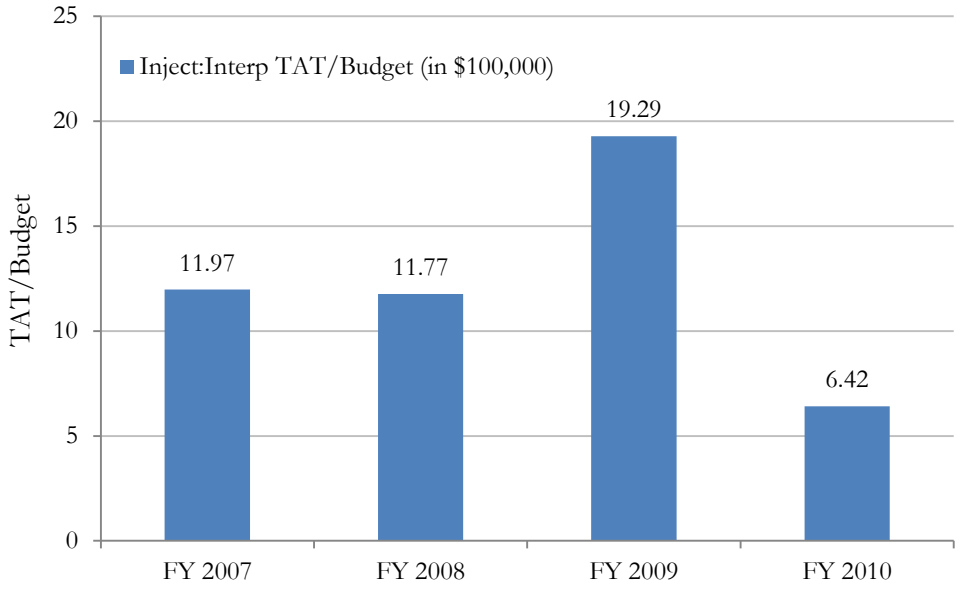


Figure 44. Stage-Level Efficiency Measure by Budget: Capillary Electrophoresis to Interpretation



Note: Because case processing data were not available for the last three months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 75 percent of the year.

Figure 45. Stage-Level Turnaround Time: Interpretation to Technical Review

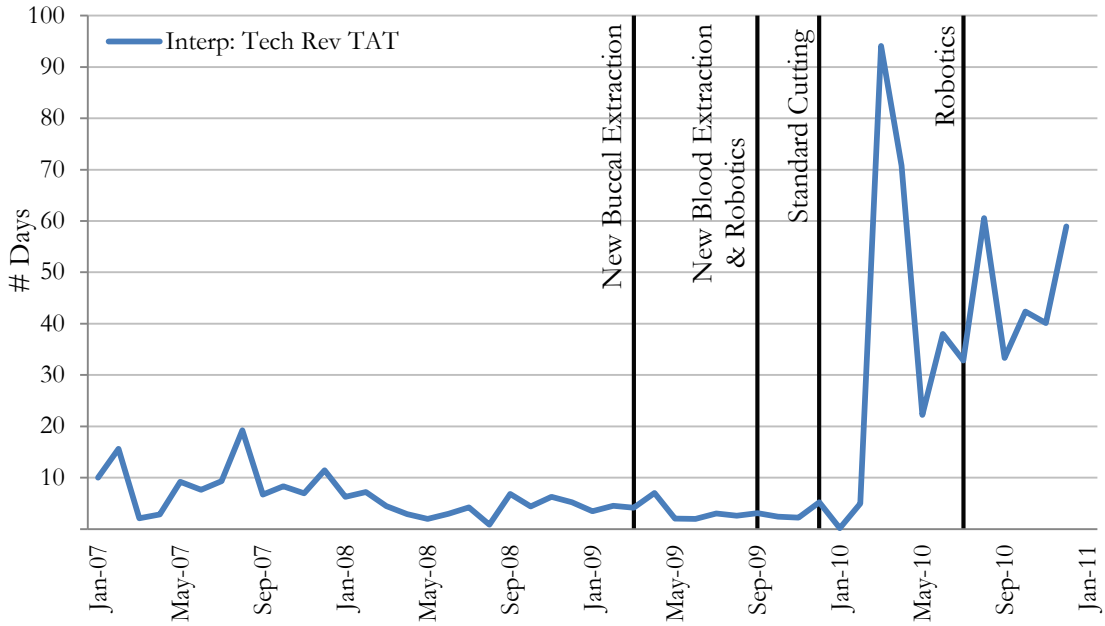


Figure 46. Stage-Level Turnaround Time: Technical Review to Report

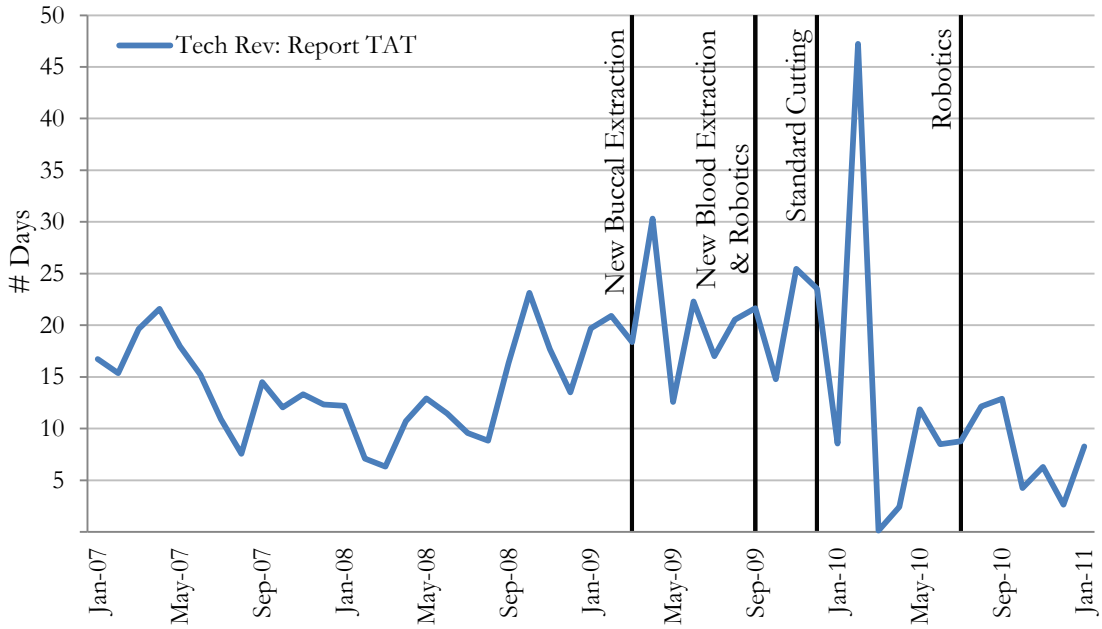


Figure 47. Stage-Level Efficiency Measure by Labor: Technical Review to Report

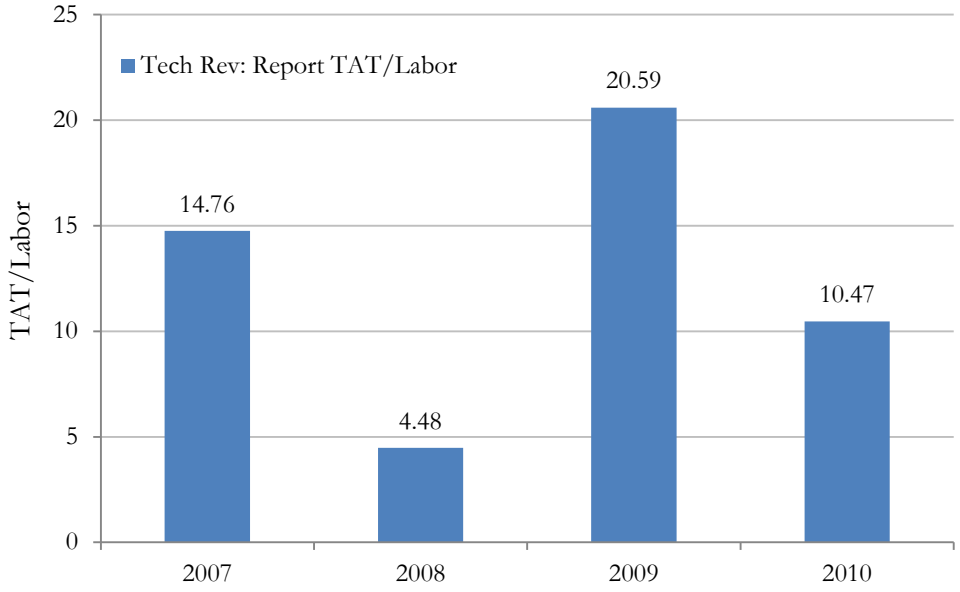
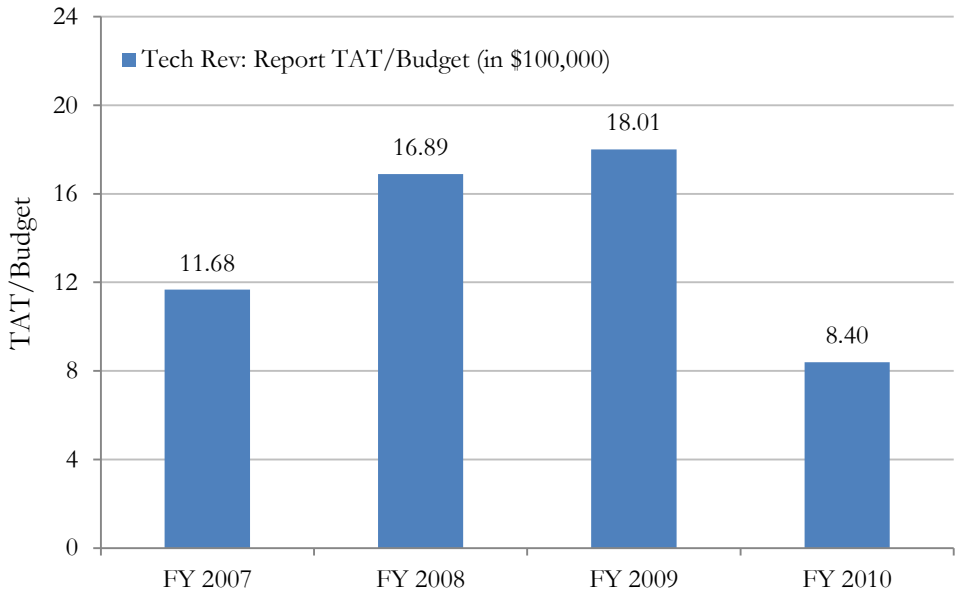


Figure 48. Stage-Level Efficiency Measure by Budget: Technical Review to Report



Note: Because case processing data were not available for the last three months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 75 percent of the year.

4.4.2 Pre/Post Throughput Comparison Tests

The research team conducted two types of tests (independent samples t-test and Mann-Whitney U test) to compare the throughput both before and after the first major implementation milestone: the use of *ZyGEM* extraction for buccal swabs. While not all grant interventions were implemented at this point in time, it marks the first point when casework had the advantage of the grant activities. In addition, all other implementations still occurred in the "post" period and should, theoretically, only strengthen the relationship as more pieces are implemented. Both the independent samples t-test and Mann-Whitney U test are used to compare differences between two groups. However, the Mann-Whitney U uses the median as the measure of central tendency and, therefore, does not require normality as the independent samples t-test does. Although t-tests are robust to normality assumption violations, the Mann-Whitney U test was also conducted since the data were skewed.

Statistical tests did not find a significant difference between the number of cases completed before and after implementation (see table 6). This result was confirmed by both tests, although the test statistic was in the expected direction with slightly higher throughput after implementation of the first grant milestone. Dividing the throughput by number of staff resulted in similar findings (although the t-test approached significance). In contrast, there was a significant difference between "pre" and "post" periods for the measure of throughput divided by budget expenditures. This result was not confirmed, however, by the Mann-Whitney U test. Further, it is unclear whether this significant finding is due to the grant itself or due merely to the fact that the entire 2010 budget period (which had much fewer expenditures than other years) happens to fall within the "post" period. Therefore, findings are inconclusive on whether the laboratory had greater efficiency in terms of financial resources due to grant activities.

Table 6. Pre/Post Comparison Tests

Throughput	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of *ZyGEM* Extraction (3/09)	-1.62	0.12	265.00	0.26
Throughput/Labor	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of *ZyGEM* Extraction (3/09)	-1.98	0.054	241.50	0.12
Throughput/Budget	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of *ZyGEM* Extraction (3/09)	-2.53	0.018	239.00	0.11

4.4.3 Regression Analyses

In order to detect whether the program had an effect on turnaround time and its related efficiency measures after controlling for other sample characteristics, the research team performed a series of negative binomial regression analyses (see table 7). Regressions were used to model overall and stage-level turnaround times, both as pure productivity measures (case turnaround time) and as efficiency measures (sample turnaround time divided by the number of staff during the year the sample began and sample turnaround time divided by the year's budgetary expenditures in $100,000 units).

The research team included a series of dummy variables to mark each major grant implementation: (1) the March 2009 implementation of *ZyGEM* extraction for buccal swab samples, (2) the fall 2009 implementation of *ZyGEM* extraction for blood samples (September) and robotics for quantification setup (October), (3) the use of standard cutting and subsequent removal of the quantification stage altogether in December 2009, and (4) the implementation of automated extraction and amplification setup with robotics in July 2010. Dummy variables were coded as a 0 or 1 depending on whether the intervention had been implemented. Regression coefficients represent the unique influence of each intervention above the effects of those interventions occurring previously. In addition, a separate event unrelated to the grant was included in the model to control for the potential effects of transitioning the LIMS in March 2010. Other non-grant events aligned with existing intervention milestones, including the October 2009 switch to *Identifiler* amplification kit and the November 2009 upgrade to a *3130* capillary electrophoresis instrument. These two events were categorized with the second and third intervention points listed above, respectively. When interpreting the effect of any variable that includes multiple milestones (grant or otherwise), there is no way to know which intervention was responsible for the change (or if the change happened to occur at the same time, but was not caused by the interventions). However, assumptions can be made about some co-occurring interventions not causing changes in stages for which they are not used (i.e., it is more likely that the *ZyGEM* extraction will cause changes during the extraction stage than the co-occurring quantification robotics which are utilized at a later stage).

Because some of the intervention milestones were in such close proximity to each other, the models exhibited multicollinearity issues when all intervention variables were included together. Multicollinearity occurs when a large proportion of variance is shared between two variables (i.e., much of the time overlaps for two or more interventions). In particular, the second and third intervention variables had tolerance and variance inflation factor scores indicative of multicollinearity. In order to prevent poor parameter estimations caused by multicollinearity, the research team only included intervention variables expected to influence each outcome of interest. The research team based these decisions on the site's recommendations and the research team's internal forensic expertise. All intervention

variables were included in the model predicting overall sample turnaround time (as opposed to stage-level turnaround times) because every intervention was expected to have some impact on turnaround time.[46]

Interventions had differing effects depending on the stage examined. For overall sample turnaround time, only the variable for implementing standard cutting and upgrading the capillary electrophoresis instrument (both occurring in late 2009) was associated with reduced processing time. Turnaround time appeared to increase for samples beginning after the implementation of *ZyGEM* extraction for buccal swabs and beginning after the LIMS transition. The stage between assignment and extraction completion showed decreases in turnaround time after the implementation of *ZyGEM* extraction (improvements did not materialize until after both the buccal and blood samples were targeted), but automated extraction was associated with an increase in turnaround time. The stage between the completions of extraction and amplification found potential effects of the grant, including decreased turnaround time related to the implementations of automated quantification, *Identifiler* kit, and standard cutting (*ZyGEM* extraction also occurred around the same time but would not be expected to have a strong effect on amplification). Again, other robotics (such as those used for amplification setup) were associated with an increase in turnaround time.

Between amplification and the completion of capillary electrophoresis, the upgraded capillary electrophoresis instrument and robotics were associated with reduced time spent during this stage. Data review time appeared to increase after changes to the extraction, quantification, and amplification kit procedures. The technical review stage experienced reduced turnaround time for two of the event dates (implementation of standard cutting/upgraded capillary electrophoresis instrument and extraction/amplification robotics), but some grant interventions (*ZyGEM* extraction for buccal samples and extraction/ amplification robotics) were associated with increased turnaround time for the report stage. The LIMS transition was associated with decreased turnaround time for all stages except the data review and report stages. A particularly large relationship was observed for the technical review stage, reasonably so since this accounts for the data anomaly caused by the change in LIMS in how technical review dates were recorded.

Other characteristics were also influential in predicting sample turnaround time. Reruns or having multiple submissions was associated with increased turnaround time on the overall processing as well as four of the six stages (two stages had negative relationships between

[46] The research team tested alternate versions of the model to determine if multicollinearity was obscuring other significant findings. Hierarchical versions of the model, models removing the third intervention/confound (standard cutting and upgraded CE instrumentation), and models with each event separately with the remaining controls did not result in substantially different results. Therefore, the full model was reported. Models with all of the event variables also had a lower AIC than alternate models.

turnaround time and sample reruns or multiple submissions).[47] Blood standards were more likely to have shorter turnaround times overall, although this finding changed by stage with half of the stages showing the opposite relationship. Analysts with more experience tended to spend more time processing DNA samples (again, possibly due to competing management responsibilities) with the exception of the interpretation and report-writing stages, where more experience was related to shorter turnaround times.

Regression findings for efficiency measures were generally similar to those of the productivity measures. Findings were identical in terms of significance and direction for the overall sample turnaround time and the same outcome divided by labor. However, when annual budget was taken into account, the overall sample turnaround time had a positive relationship with automated extraction and amplification and a negative relationship with the second intervention milestone (*ZyGEM* extraction for blood samples, automated quantification, and the non-grant-related switch to the *Identifiler* kit); standard cutting and use of the *3130* instrument were no longer significant.

Stage-level efficiency measures also generally aligned with their respective productivity measures, although there were some changes when taking into account annual labor or budgetary expenditures. For instance, the initial *ZyGEM* extraction of buccal swabs became significant (in a positive direction indicating increased turnaround time) once labor was taken into account for the stage between assignment and extraction completion. The second intervention/confound milestone became significant once budget was accounted for, and automated extraction and amplification was no longer associated with increased turnaround time for turnaround time divided by annual labor counts (although the *p*-value was only slightly greater than 0.05 for this model).

Overall, the grant program appeared to be related to changes in the turnaround time of known standards after accounting for other events and factors. However, these influences varied by stage and in direction, resulting in an unclear picture of the true impacts of the grant on turnaround time. The absence of substantial change in overall turnaround time may be due to (a) weak effects of the grant activities, (b) the loss of specific time-savings during other stages of the process, or (c) competing gains and losses in turnaround time from different events and interventions. Further, the incremental nature of implementation, paired with milestones occurring in close proximity to each other (thus, "muddying" the waters), made detecting and interpreting changes difficult.

[47] While multiple submissions and reruns may mask the turnaround time of discrete routings, regression analyses control for this and allow researchers to understand the unique influence of other variables while controlling for whether a sample was rerun or was part of a case with multiple submissions.

Table 7. Regression Results for Kansas City

Overall TAT Regression	Overall Sample TAT		Overall Sample TAT/Labor		Overall Sample TAT/Budget	
	b coeff.	*p*	*b coeff.*	*p*	*b coeff.*	*p*
Intervention: *ZyGEM* Extraction for Buccal	0.22	<.01	0.37	<.01	0.11	0.02
Intervention: *ZyGEM* Extraction for Blood; Quant Robotics; Confound: *Identifiler* Kit	-0.14	0.06	-0.14	0.07	-0.18	0.03
Intervention: Standard Cutting/Stop Quant; Confound: Upgraded CE	-0.17	0.03	-0.27	<.01	-0.16	0.06
Intervention: Extract/Amp Robotics	-0.01	0.73	-0.02	0.61	0.14	<.01
Confound: LIMS Transition	0.24	<.01	0.19	<.01	0.85	<.01
Rerun or Multiple Submissions	0.72	<.01	0.72	<.01	0.71	<.01
Blood Standard	-0.17	<.01	-0.18	<.01	-0.12	<.01
Analyst Experience	0.01	0.03	0.01	0.02	0.03	<.01

Stage-Level Regression	Assign–Extract TAT		Extract–Amp TAT		Amp–CE Inject TAT		CE Inject–Interpret TAT		Interpret–Tech Review TAT		Tech Review–Report TAT	
	b coeff.	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*
Intervention: *ZyGEM* Extraction for Buccal	0.05	0.52	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.35	<.01
Intervention: *ZyGEM* Extraction for Blood; Quant Robotics; Confound: *Identifiler* Kit	-0.37	<.01	-0.46	<.01	-0.01	0.97	0.73	<.01	-0.30	0.12	-0.18	0.12
Intervention: Standard Cutting/Stop Quant; Confound: Upgraded CE	N/A	N/A	-0.34	<.01	-0.55	<.01	0.34	0.07	-1.06	<.01	0.17	0.19

Stage-Level Regression	Assign–Extract TAT		Extract–Amp TAT		Amp–CE Inject TAT		CE Inject–Interpret TAT		Interpret–Tech Review TAT		Tech Review–Report TAT	
Intervention: Extract/Amp Robotics	0.81	<.01	3.09	<.01	-1.31	<.01	N/A	N/A	-0.42	<.01	0.23	0.01
Confound: LIMS Transition	-0.40	<.01	-0.88	<.01	1.89	<.01	-2.09	<.01	3.55	<.01	-0.89	<.01
Rerun or Multiple Submissions	1.53	<.01	-0.17	<.01	0.18	<.01	0.38	<.01	0.35	<.01	-0.21	<.01
Blood Standard	-0.41	<.01	0.01	0.84	-0.21	<.01	-0.07	0.32	0.21	0.01	0.04	0.43
Analyst Experience	0.04	<.01	0.07	<.01	0.03	0.01	-0.16	<.01	0.04	<.01	-0.05	<.01

Stage-Level Regression	Assign–Extract TAT/Labor		Extract–Amp TAT/Labor		Amp–CE Inject TAT/Labor		CE Inject–Interpret TAT/Labor		Interpret–Tech Review TAT/Labor		Tech Review–Report TAT/Labor	
	b coeff.	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*
Intervention: *ZyGEM* Extraction for Buccal	0.21	0.01	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.62	<.01
Intervention: *ZyGEM* Extraction for Blood; Quant Robotics; Confound: *Identifiler* Kit	-0.41	<.01	-0.31	0.01	0.01	0.96	0.83	<.01	-0.15	0.40	-0.17	0.12
Intervention: Standard Cutting/Stop Quant; Confound: Upgraded CE Instrument	N/A	N/A	-0.44	0.01	-0.61	0.05	0.26	0.12	-1.19	<.01	0.12	0.36
Intervention: Extract/Amp Robotics	0.76	<.01	1.80	<.01	-1.29	<.01	N/A	N/A	-0.42	<.01	0.18	0.05
Confound: LIMS Transition	-0.52	<.01	-1.41	<.01	1.97	<.01	-1.98	<.01	3.43	<.01	-1.34	<.01
Rerun or Multiple Submissions	1.53	<.01	-0.09	0.07	0.17	0.05	0.47	<.01	0.50	<.01	-0.29	<.01
Blood Standard	-0.48	<.01	0.06	0.21	-0.36	<.01	-0.03	0.65	0.15	0.02	-0.02	0.65
Analyst Experience	0.04	<.01	0.05	<.01	0.08	<.01	-0.14	<.01	0.05	<.01	-0.05	<.01

Stage-Level Regression	Assign–Extract TAT/Budget		Extract–Amp TAT/Budget		Amp–CE Inject TAT/Budget		CE Inject–Interpret TAT/Budget		Interpret–Tech Review TAT/Budget		Tech Review–Report TAT/Budget	
	b coeff.	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*	*b coeff.*	*p*
Intervention: *ZyGEM* Extraction for Buccal	-0.03	0.73	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.28	<.01
Intervention: *ZyGEM* Extraction for Blood; Quant Robotics; Confound: *Identifiler* Kit	-0.39	<.01	-0.59	<.01	-0.15	0.43	0.50	<.01	-0.49	0.01	-0.22	0.17
Intervention: Standard Cutting/Stop Quant; Confound: Upgraded CE Instrument	N/A	N/A	-0.34	0.01	-0.54	0.02	0.38	0.05	-1.03	<.01	0.18	0.32
Intervention: Extract/Amp Robotics	0.90	<.01	3.27	<.01	-1.20	<.01	N/A	N/A	-0.22	0.04	0.34	<.01
Confound: LIMS Transition	0.23	0.02	-0.27	0.03	2.58	<.01	-1.29	<.01	4.04	<.01	-0.65	<.01
Rerun or Multiple Submissions	1.53	<.01	-0.19	<.01	0.23	<.01	0.39	<.01	0.51	<.01	-0.35	<.01
Blood Standard	-0.38	<.01	0.09	0.02	-0.18	0.01	-0.07	0.35	0.27	<.01	0.03	0.64
Analyst Experience	0.05	<.01	0.10	<.01	0.07	<.01	-0.14	<.01	0.07	<.01	-0.03	<.01

4.4.4 Conclusions

The Kansas City Police Crime Laboratory proposed to create a more streamlined workflow for processing known standard evidence samples through a new sample preparation and extraction technique, automation with robotics, and an expert system. The lab successfully completed all goals of their grant proposal and did not encounter substantial implementation challenges with the exception of (1) delays in obtaining approval from the NDIS Board for use of the expert system and in obtaining a working heat block for robotics, and (2) difficulty prioritizing research and development tasks while continuing normal casework responsibilities.

Across time, four intervention implementation milestones were examined. Unfortunately, the implementation of the expert system occurred too late in the study period to measure its effects. While monthly trends showed an increase in throughput after the implementation of robotics, pre/post comparison tests had mixed findings for the impact of the grant program of number of cases completed. When financial resources were accounted for, effects were stronger.

There did not appear to be a strong effect on overall sample turnaround time. Month-to-month figures did not reveal a clear pattern of change after implementation, and regression analyses showed conflicting effects for the grant activities (some were associated with longer while others were associated with shorter turnaround times). Stage-level turnaround times showed different trends from the overall sample turnaround time, although large variability made it difficult to detect effects. Disruption caused by a LIMS transition also complicated data patterns. While regression analyses provided support for beneficial effects of various components of the grant program, these improvements were often lost in the overall turnaround time or balanced out by other intervention components related to *increased* turnaround times. When accounting for labor and financial resources, 2008 and 2009 were typically the least efficient years for turnaround times, although findings varied somewhat by processing stage.

Other factors were also related to turnaround times. Samples experiencing reruns or belonging to cases with multiple submissions tended to take longer to process. The type of known standard (buccal vs. blood) had conflicting relationships with turnaround time, depending on the stage. Finally, analysts with more experience tended to spend more time processing DNA samples with the exception of the interpretation and report-writing stages, when more experience was related to shorter turnaround times.

In conclusion, Kansas City had substantial success in terms of developing, validating, and implementing all of its grant proposal components. The lab created a new workflow for known standards that more closely matched the processing of convicted offender and arrestee samples. Unlike other sites, Kansas City had few implementation roadblocks and was able to

implement all project components (the majority of them implemented at an early stage). The lab felt confident that these changes had strong impacts on the throughput and turnaround time of known standards, although the current analyses were more equivocal. Unfortunately, the research team could not measure the throughput and turnaround time outcomes of the expert system, which was implemented too late in the study period (due to a lengthy NDIS approval process) to be included in analyses. It is possible that the fully implemented workflow—with expert system—could have tipped the productivity and efficiency gains to a level detectable with statistical tests. However, current analyses could not identify a consistently significant effect for samples processed through January 2011.

5. CASE STUDY: LOUISIANA STATE POLICE CRIME LABORATORY

5.1 Overview of the Laboratory

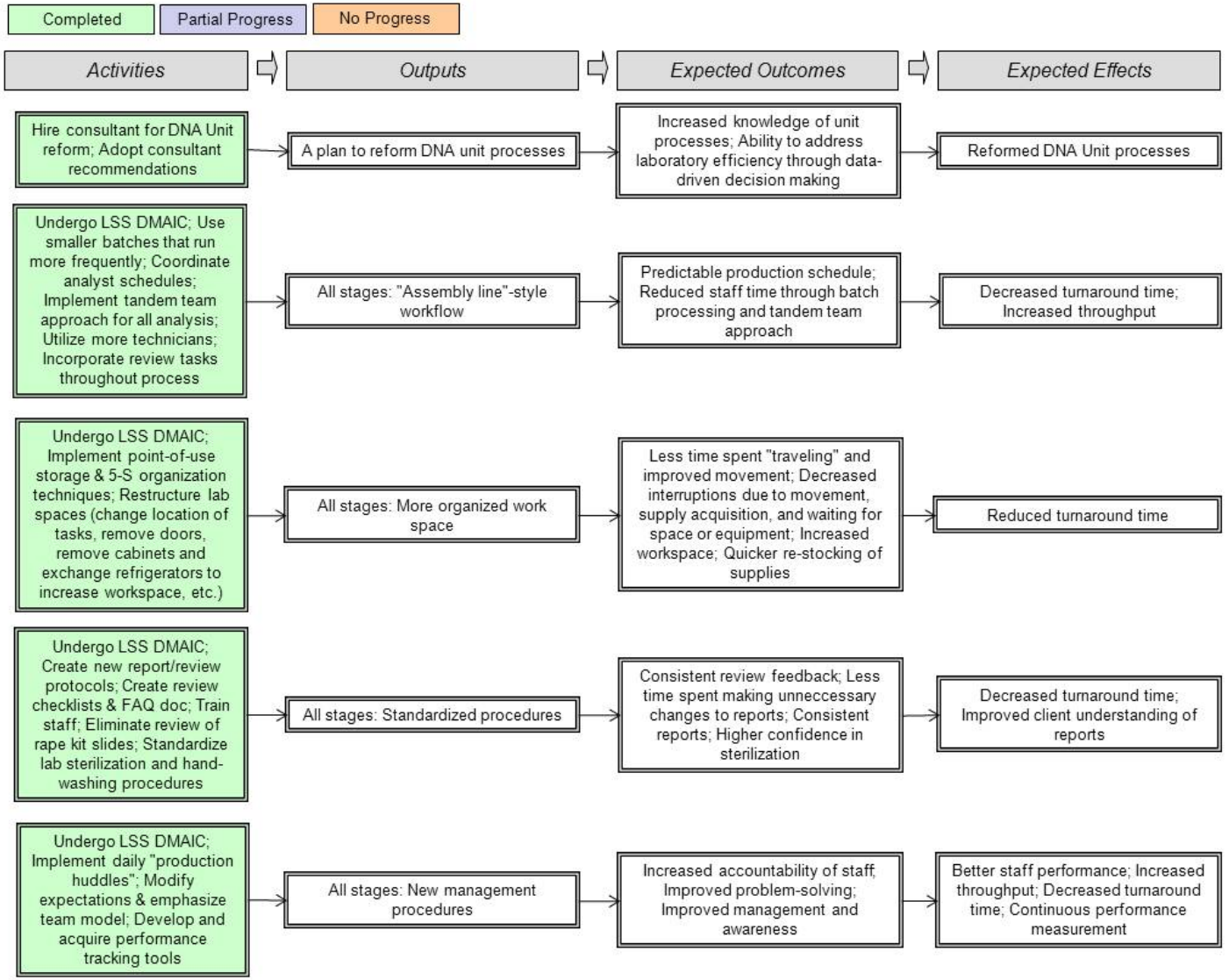
The Louisiana State Police Crime Lab (hereafter, Louisiana or LSP) is an accredited public crime laboratory located in Baton Rouge. The laboratory accepts forensic evidence from casework, offender, arrestee, and missing person DNA samples from the state of Louisiana (although the majority of their work is from the Baton Rouge area).

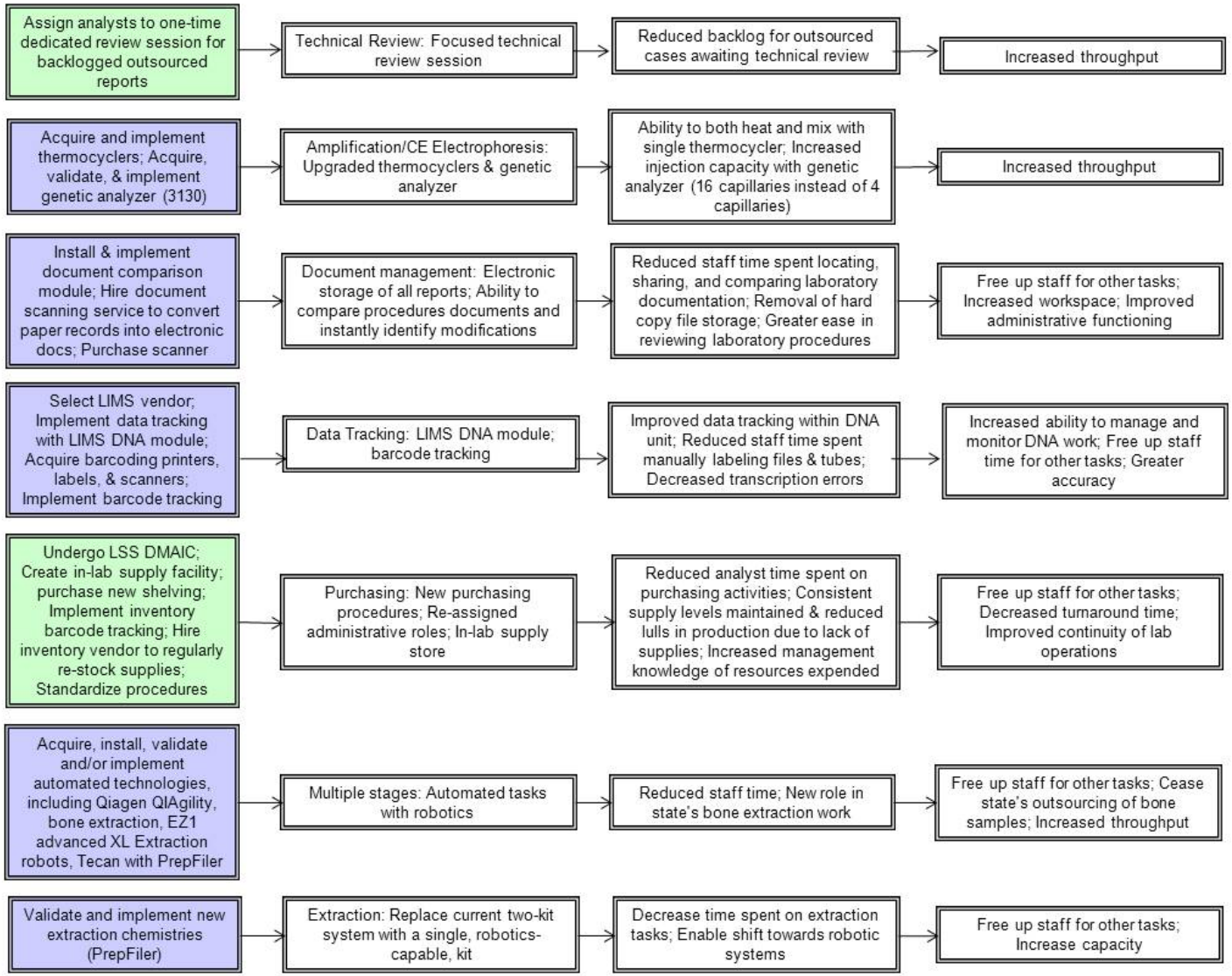
Louisiana proposed to hire a consultant to conduct process mapping and make recommendations based on this mapping to improve efficiency within the lab. Their proposal also included acquiring and/or validating a number of new technologies that would improve their processes, but the acquisition of these instruments was at least partially dependent on the recommendations received from the process mapping vendor.

5.2 Description of Grant Goals

NIJ awarded Louisiana a $450,000 Forensic DNA Unit Efficiency Improvement Program grant to fund the purchase of equipment, contracting with a process improvement vendor, contracting with a LIMS vendor, software licenses, and on-site training. Louisiana's 25 percent nonfederal match supported contracts with the LIMS vendor and the consultant for process mapping, plan development, and implementation support. The proposal did not describe from the outset which specific tasks would be accomplished, instead stating Louisiana would hire a consultant to help determine the best ways to improve efficiency. By the end of the study period, Louisiana had conducted many changes to the DNA process and general administrative procedures (see figure 49). Three main goals of these interventions were to (1) improve DNA Unit analysis capacity and productivity, (2) leverage technology to increase efficiency, and (3) sustain established improvements with clerical time savers. Louisiana predicted that the grant activities would create a 50 percent decrease in case backlog.

Figure 49. Logic Model for Louisiana





Louisiana's grant activities fall into four main categories: (1) Lean Six Sigma reforms for DNA casework, (2) new technologies and chemistries, (3) document and data management improvements, and (4) Lean Six Sigma reforms for purchasing procedures. Each is described below. Louisiana first hired a consultant to implement Lean Six Sigma reforms for the DNA workflow. Lean Six Sigma is a process improvement method that seeks to eliminate waste, improve efficiency, and increase productivity while improving quality. It is a hybrid of *Lean Thinking* and *Six Sigma* methods. *Lean Thinking* evolved from the production-line activities first implemented in the automotive industry. Its goal is to reduce the time from customer request to the final deliverable. This is accomplished by eliminating activities that do not add value, from the customer's point of view, to the final product. *Six Sigma* is a management strategy that seeks to improve processes through an accurate understanding of industry processes and abilities, data-driven decision making, and sustainable actions. When combined into a Lean Six Sigma (LSS) approach, this hybrid works from seven core principles: "focus on the customer, identify and understand how the work gets done, manage, improve and smooth process flow, remove non-value added steps and waste, manage by fact and reduce variation, involve and equip the people in the process, and undertake improvement activity in a systematic way" (Richard and Kupferschmid 2011).

The LSS framework for creating change includes a set of five stages: Define, Measure, Analyze, Improve, and Control (DMAIC). For the Louisiana lab, each of these stages involved a series of consultant-guided exercises and activities to better understand the lab's current functioning and areas for improvement. During the Define stage, the project was outlined (scope, goals, clients, stakeholders, supplies, etc.), a project charter was developed, and a process map was created. During the Measure stage, current practices were tracked and measured. Current practice performance metric data, collected in the Measure phase, were analyzed in the Analyze stage. During the Improve phase, team members designed and piloted a new process that directly addressed the bottlenecks identified in the Measure and Analyze phases. Finally, during the Control stage, management worked to sustain the increased level of productivity. The site expected that these improvements would increase case processing productivity, as well as improve management and the lab's relationships with submitting agencies.

The newly developed process included many changes to the lab. The lab made organizational and structural changes to the workspace, relying on the principles of point-of-use-storage (storing all necessary supplies and equipment in the area where related procedures are performed) and "5-S" (a systematic method to maintain a neat and clean work area with specified and labeled locations for every item). The lab also standardized procedures from sanitation techniques to evidence screening procedures to the protocol for technical reviews. Finally, the largest change was the coordination of a new "assembly line" style workflow. The new schedule carefully coordinated analysts to have specific

responsibilities for each day of the week with a continuous and integrated workflow. The new schedule was expected to result in at least eight cases completed per day. Analysts were assigned to weekly teams with each analyst responsible for a different part of analysis, a change from the existing philosophy where each analyst is responsible for his or her own case. Part of this change was to combine samples into smaller batches that could be run more routinely. In order to prepare for the large change in workflow, the lab also participated in a focused technical review session, which lasted several days. This focused session was designed to substantially reduce the backlog of outsourced case reports awaiting secondary review.

Across all of these LSS activities, a new management approach was implemented: one that emphasized continuous monitoring, problem-solving, and accountability. A large component of this new management approach was the institution of daily meetings called "production huddles" where the staff assigned and planned casework, engaged in problem-solving for work challenges, and compared actual performance with laboratory goals. An organized board was created to track assignments and progress of analysts; later, this board was replaced with an electronic version (the *i-dashboard*) capable of more sophisticated features.

Louisiana planned to implement a number of new technologies to improve DNA processing efficiency. Upgraded thermocyclers and a new genetic analyzer with increased injection capacity, if implemented, would provide the lab with more sophisticated instruments. The lab also included new extraction robotics and bone extraction equipment in its plan to improve the extraction processes and add a new capability for the state's forensic services (bone extraction was typically outsourced). Through a previous grant, Louisiana had acquired three *Qiagen QIAgility* robotic systems, but prior to implementation into casework these systems needed to be validated. Louisiana proposed to outsource this validation, along with the validation of a new extraction chemistry, *PrepFiler*, to *Applied Biosystems*. Outsourcing validation tasks enabled the laboratory to keep all DNA analysts actively working on casework during the validation period.

Louisiana also proposed to improve data tracking and documentation management. The lab wanted to purchase a DNA-specific LIMS module and expand upon their already-existing barcoding project (funded through other means) with additional barcoding equipment. Implementing barcoding technologies to track case-level information, sample-level identification information could help to eliminate staff time spent on manual creation of labels and reduce transcription errors. Additionally, the site proposed to improve document storage and comparison by undertaking an effort to convert over 310,000 records to electronic formats and to integrate a document comparison module into their existing LIMS (for the purpose of policy and procedures review and updating).

The final component of Louisiana's efficiency strategy was to reduce the clerical duties performed by DNA analysts. The site proposed to perform a second LSS reform through an additional consultancy targeted on the laboratory purchasing department. At the beginning of the grant period, DNA analysts were heavily involved in the purchasing of supplies, for both their laboratory work as well as office supplies. The site estimated that laboratory staff members were making 200 trips to office supply stores in a single six-month period. It was anticipated that shifting these non-analysis responsibilities to clerical staff would increase the amount of time each analyst had for DNA processing.

5.3 Implementation Findings

5.3.1 Implementation Description

Figure 50. Louisiana Implementation Timeline October 2008–December 2010

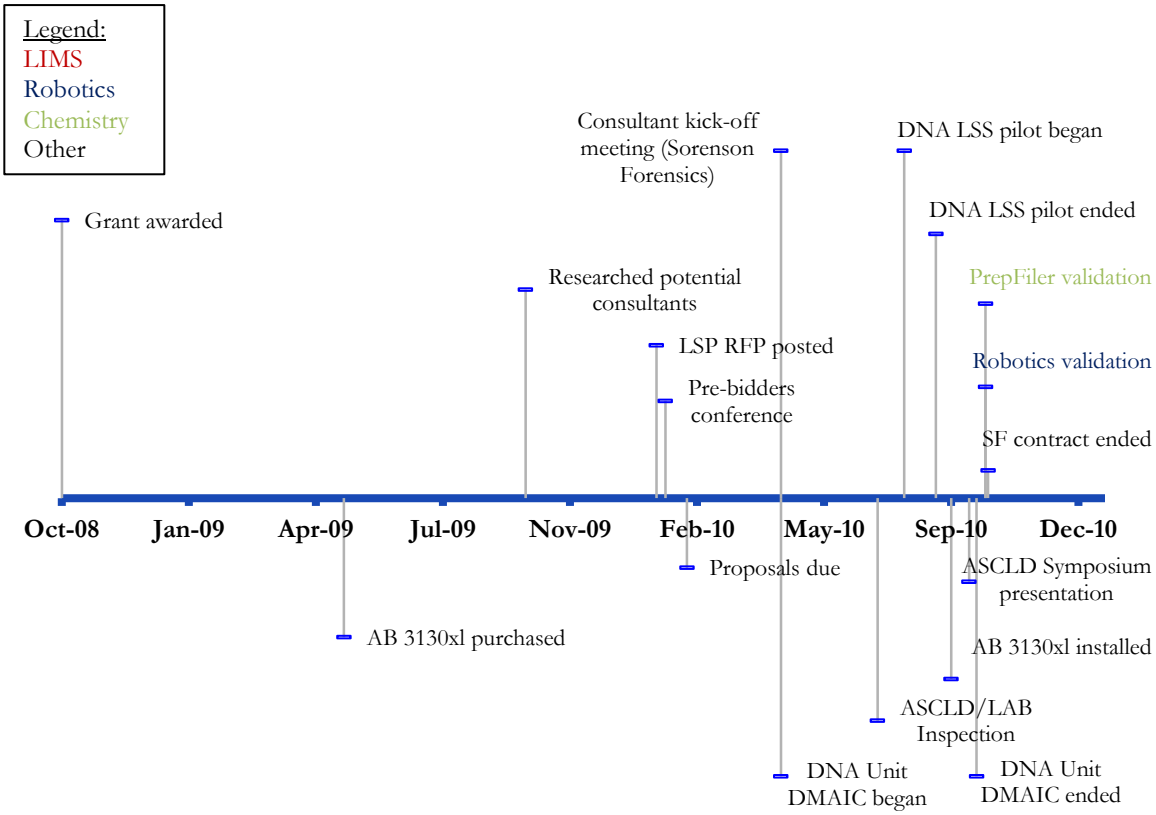
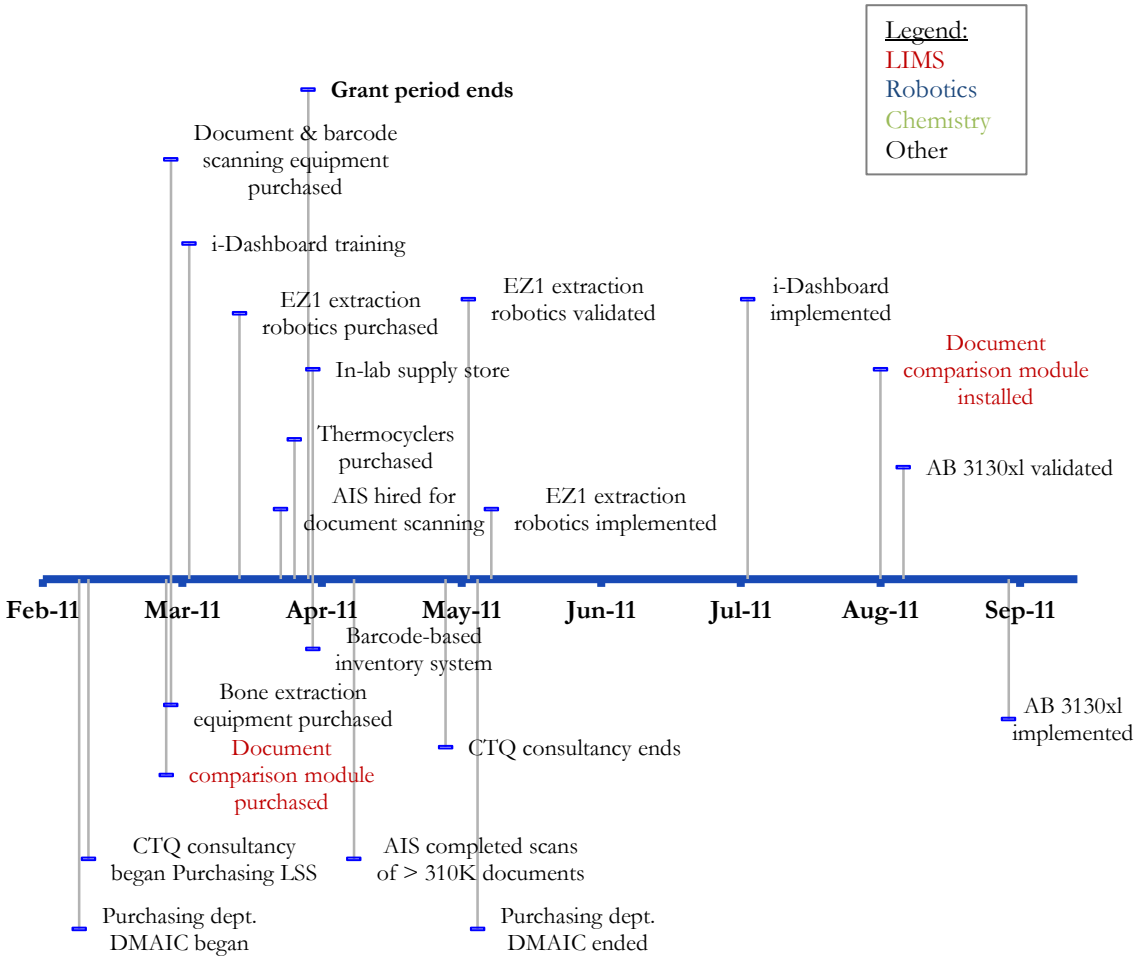


Figure 51. Louisiana Implementation Timeline February 2011 – September 2011



Louisiana concluded their grant period on 3/31/2011, 29 months after notification of their award. They received several project extensions to support a number of implementation challenges, which are discussed below.

The consultant request for proposals (RFP) was posted on January 12, 2010, and LSP hosted a pre-bidders conference for seven companies. The submission from Sorenson Forensics was selected as the winning proposal in April that year, and they were awarded a contract for $174,950. The consultancy began that month with a three-day kick-off meeting at the laboratory. During the project period, the LSP team met with the consultants for a total of 29 days between April 20, 2010, and September 30, 2010.

As noted above, the main methodology that the Sorensen/LSP team adopted was Lean Six Sigma, which includes a series of stages: Define, Measure, Analyze, Improve and Control.[48] The DMAIC process began in April 2010. The Define stage resulted in (1) the creation of a project charter, (2) a description of current practices through process mapping, and (3) maps that trace the physical movement of case evidence through all processing stages. This stage was completed in May 2010.The Measure phase involved the collection of data on the activities of the process in order to determine the baseline level of performance. The team reviewed their activities at each processing stage. This exercise revealed that 97.3 percent of the 186-day turnaround time (from case assignment to report release) was "non-value added" time.[49] The team also identified the root cause of delays and the resources needed to accomplish DNA processing goals, among other activities. The Measure stage was completed by mid-May 2011. The Analyze phase used the data gathered during previous stages to identify the location and causes of process bottlenecks. Improved processes were then designed so that bottlenecks were reduced and the workload in the lab was consistent, with little variation and few spikes or lulls.

During the Improve phase of the DMAIC process, LSP conducted a four-week pilot of their new procedure. The new process was a five-day cycle where three DNA analysts were supported by four technicians. The DNA unit had a staggered adoption; one team began the new system on the first week, a second team joined on the second week and by the third week, all three teams were participating in the new process. In preparation for the new staggered task schedule, improvements were made to the lab and analyst workstations, and excess motion or walking was reduced by relocating equipment, repurposing individual laboratory rooms' lab processes, and removing lab doors. Workstation setups were standardized, and manual tube labeling was replaced with new barcode labeling methods (the lab used grant funds to supplement an already-existing project to institute barcode tracking at the lab).

The final DMAIC phase, the Control phase, ensured that methods were in place to continually collect staff performance metric data so that management could ensure that the new productivity levels were maintained. This phase was dependent on workload data to make informed management decisions. The Control phase began in early September 2010 and continued until the end of May 2010. Managers and staff continued the piloted changes and used real-time performance measures to make informed decisions on a daily basis.

[48] For more in-depth information, please see site's final technical report to NIJ.

[49] Non-value added time refers to time where there is no actual processing tasks being performed. From the customer's point of view, this would include the time evidence spends sitting in the property room or waiting to be assigned to an analyst.

At the completion of the consulting work, the project team gave a presentation to local stakeholders and command staff that described what changes they made and the results of these change. The site reported that this presentation was well received and was one of the key steps in maintaining communication with command and establishing understanding of appreciation for the project interventions. During the last month of the consultancy, Sorenson Forensics presented the approach and preliminary outcomes at the 2010 ASCLD Symposium.[50]

At the outset of the DMAIC Control phase, the performance data needed for management was accomplished by hand calculations and displayed on a whiteboard. To reduce the time spent producing these metrics, LSP staff acquired an electronic display and i-dashboard software that interfaced with their LIMS and automatically calculated performance metric data. By the end of the project period, the display outputs had been designed but the system was not yet implemented. The site reported that these graphics will be used to inform daily DNA "production huddle" meetings. Once fully implemented, LSP expects this technology to provide staff and managers with the data necessary to identify and resolve problems.

In addition to the work conducted in the DNA unit, LSP applied Lean Six Sigma techniques to the purchasing and procurement process at the laboratory. Before these reforms, DNA analysts were partially responsible for purchasing supplies, managing inventory, managing budgets, and communicating with vendors. LSP used grant resources to hire CTQ Consultants to establish a LSS system for the purchasing department. CTQ Consultants was awarded a contract for $16,000 to apply LSS processes to the LSP Business Unit purchasing activities. The consultants and a six-member LSP team used the DMAIC process to institute reforms. The purchasing DMAIC review took place from February through May 2011. During that time, the average purchasing time declined from 40 business days to 7, according to the lab. LSP staff estimate that before the implementation of this system, laboratory staff were making 200 trips to office supply stores in a six-month period. Now, supplies are marked with barcodes that are scanned when taken for use in the lab. An inventory management vendor replenishes expended supplies on a monthly basis. This change in purchasing procedures has shifted many clerical tasks from the laboratory analysts to administrative staff, resulting in an increase in time that analysts can devote to casework.

To aid in document management, LSP hired a professional service, Advanced Imaging Solutions, to scan more than 310,000 printed documents, including quality control records for outsourced cases, records of training, instrument maintenance and validation, DNA analysis stage worksheets, and CODIS documentation. These paper documents were

[50] T. D. Kupferschmid, "100% Increase in Laboratory Productivity due to Implementation of Lean-Six Sigma Practices," ASCLD Symposium, Baltimore, MD, September 15, 2010.

removed from the lab workspace, clearing 600 feet of shelving and eight file cabinets. In order to sustain this effort, Louisiana purchased a high-speed scanner for clerical staff use.

LSP had previously acquired Qualtrax software through an earlier grant. Funds from this grant award were used to purchase a document comparison module for this system. This module can electronically compare documents and flag changes. Once implemented, this technology will assist with version management of procedural documents. In addition to the Lean Six Sigma consultancy work, grant resources were used to purchase several technologies designed to expand workstations, reduce administrative burden on lab analysts, and limit manual sample handling. New thermocyclers were purchased but not implemented by the end of the evaluation period. Workstations were outfitted with barcode scanners, barcode printers, and document scanners. These tools aided the integration of barcoded case files and sample tracking in the laboratory.

Two *Qiagen EZ 1Advanced xL* extraction robotic units were purchased but not implemented during the grant period. These systems were designed to reduce manual sample handling by lab analysts and to increase the overall automation of DNA processing. LSP purchased and validated the Applied Biosystems *Prepfiler* extraction kit. They proposed to use this kit for extractions because it is compatible with the two robotic systems, *Tecan* and *QIAgility*. Validation was completed for two types of reference sample types, buccal swabs and bloodcards (non-FTA)[51]. During the project period the kit was not implemented into casework procedures. Implementation was expected by July 2011.

LSP selected Applied Biosystems (AB) to supply the robotics validations. Three *Qiagen QIAgility* robotic systems, one for quantification setup, one for amplification setup and one for separation setup, were acquired with funds from other grants, and the validations of two of these units was funded through the Unit Efficiency grant. AB completed the validation process and LSP staff training November 2010. However, by the end of the grant period, the validated robotics were not implemented into laboratory casework. The genetic analyzer was purchased and installed but by the end of the grant period, the validation was ongoing.

Finally, the last new technology acquired was equipment to perform DNA extraction on bone samples. Before the acquisition of this equipment, LSP had to outsource all bone evidence. The extraction equipment was not installed or implemented during the grant period. However, the site expects that once implementation is achieved, they will be able to end outsourcing of bone items.

[51] Non-FTA cards contain an absorbent paper that is designed for the short term-storage of blood or bodily fluids. FTA cards are also used for body fluid storage, but they contain chemicals that help protect and preserve DNA for long-term storage.

5.3.2 Implementation Challenges

As stated in the Implementation description, Louisiana was delayed in the initiation of their project. The primary reason for this delay was the competing obligations from other grants. Ultimately, the lab was able to successfully implement major reforms to both the DNA unit and lab purchasing processes within a relatively short period of time once grant activities began. This is illustrated by the timelines shown in figures 50–51.

During the Lean Six Sigma DMAIC process, the DNA unit staff were tasked with several time-consuming take-home assignments. While these did not cause any implementation delays, the site reported that it was an additional burden that made it difficult to maintain other casework obligations. The project manager also reported that bureaucracy-driven delays in purchasing were a challenge to implementation.

Many of the technologies purchased with this grant were not implemented during the grant period of performance. LSP staff had considered purchasing a DNA module for their existing LIMS , a *JusticeTrax* product. Ultimately, LSP did not acquire a DNA module because they could not find a product with the specifications they desired at their budgeted price. The grant funds dedicated for the DNA module were repurposed to support additional LSS activities, instrument validations, the document scanning project, and the purchase of additional equipment. LSP reports that a decision was made to prioritize the adoption of the new LSS system for the DNA unit rather that install, validate, and implement the purchased robotics. This decision was due in part to their preparations for ISO 17025 laboratory accreditation.[52] Most grant activities ceased during the preparation for and during the inspection itself. As a result, implementation of some purchased technology did not occur within the evaluation period.

5.3.3 Final Perceptions

At the end of the study period, the key contact at Louisiana reported that the interventions implemented because of the 2008 DNA Unit Efficiency grant had "changed their world." The lab felt the grant had substantially impacted its ability to conduct DNA casework and reported that they were no longer outsourcing DNA samples. The Lean Six Sigma reforms in both the DNA unit and the purchasing unit also spread to additional sections of the LSP laboratory, while other laboratories and the state government have also shown interest in learning about its implementation. The interviewee reported that beyond the improvements in throughput, the LSS intervention made it possible for the lab to have greater control and predictability over its casework (and to be able to communicate more accurate expectations to customers), as well as have time for other pursuits such as research and validation studies.

[52] International Organization for Standardization number 17025 is a standard used to accredit testing and calibration laboratories. Preparation for accreditation can be very time consuming.

The project manager thought it was an important lesson learned that significant changes could occur without equipment purchases or major changes to actual methods and instead could be influenced by scheduling and organizational changes. While some new equipment was purchased for the LSS reforms, this equipment did not radically change the procedures they used to process DNA. The site expressed that the cost of implementing the LSS process was small for the impact that it made on their laboratory (additional changes made later, such as equipment purchases or the scanning project, added additional costs). The biggest challenges the interviewee reported were the late start due to needing to wrap up other grants, the relatively short period to produce an RFP document (two months), difficulty incorporating LSS exercises into work schedules with competing casework responsibilities, and initial resistance on the part of some staff. While some staff were hesitant to use the new DNA process in the beginning, all staff were reportedly highly satisfied with the process by the end of the study. In fact, one mentioned benefit of the LSS intervention was increased morale. The interviewee said, in retrospect, it may have been advantageous to hire someone to assist with the LSS exercises to prevent overloading other caseworking staff.

Louisiana reported a positive working relationship with NIJ, saying they were very open to Louisiana's goals. The lab also reported a continued commitment to implementing the remaining pieces of the grant, including the i-dashboard, document comparison module, and other new technologies currently undergoing validation (e.g., bone extraction equipment). The lab is also currently waiting for additional vendor developments in DNA modules before attempting this purchase.

5.4 Outcome Findings

The following section describes the data used to assess the outcomes of the NIJ Forensic DNA Unit Efficiency Improvement Program on Louisiana and changes in productivity and efficiency at Louisiana.

5.4.1 Descriptive Statistics and Trend Analysis

During the evaluation period, Louisiana had 2,748 non-outsourced forensic DNA cases,[53] 39 percent of which were canceled at some point.[54] Cases had a range of 1–214 items with an average of 7.91 items per case. The majority (72.9 percent) of cases had at least one known standard. One-quarter (24.8 percent) of cases had a sample with a high likelihood of containing a male-female mixture that may have required differential extraction, such as

[53] Cases exclude convicted offender or arrestee samples.

[54] Cases are cancelled when the submitting agency says that DNA testing is unnecessary after submitting evidence.

intimate swabs from a sexual assault kit, condom, or swabs from other body parts.[55] A smaller number (13.7 percent) of cases had textile evidence such as clothing or bedding objects, and nearly half (48.4 percent) of cases had some other type of swab other than those categorized under known standards or intimate samples. There were 35 individual DNA analysts assigned to these cases.

The majority of cases (53.9 percent) were for violent crimes, particularly homicide (21.0 percent) and sexual assault (22.5 percent). Nearly one-third (32.0 percent) of cases were for property offenses, and only a small number (1.2 percent) were drug related. Other types of offenses made up the remaining 11.6 percent of cases.

During the study period, the median throughput was 17 non-outsourced cases per month (see table 8). The median turnaround time for a case (from assignment to administrative review) was 48 days. The median time from assignment to report was 32 days. In contrast, the median time between report and technical review was 0 days and between technical and administrative reviews was 1 day. Nearly a month (median 28 days) passed between when the request was made to the laboratory and the case was first assigned.

Statistics for efficiency indices (throughput and turnaround time divided by annual labor counts and budget expenditure estimates [in $100,000 units]) are also shown in table 8. When accounting for staff resources, the lab completed about 1.45 cases per month per analyst.

[55] A differential extraction will separate sperm DNA from epithelial cell DNA. This technique is more time and labor intensive that regular extraction techniques.

Table 8. Louisiana Throughput and Turnaround Time Outcomes

Non-outsourced DNA Samples (N = 1,691)		Productivity/Labor	Productivity/Budget	Cleaned Productivity	Raw Productivity
Overall Outcomes					
Case Turnaround Time	Mean	6.01	5.18	79.40	76.27
Assignment–Admin Review	Median	3.14	2.65	48.00	36.00
	Std. Dev.	6.85	5.86	80.01	111.50
	Range	(0, 36.77)	(0, 37.47)	(0, 393)	(-48, 1121)
Case Throughput	Mean	2.09	1.87	N/A	30.49
	Median	1.45	1.34	N/A	17.00
	Std. Dev.	1.84	1.60	N/A	31.10
	Range	(0.09, 6.99)	(0.07, 6.08)	N/A	(1, 116)
Stage-Level Turnaround Time					
Request–Assignment	Mean	6.40	6.39	91.92	212.07
	Median	1.91	1.81	28.00	125.00
	Std. Dev.	10.28	10.69	152.77	246.10
	Range	(0, 88.06)	(0, 80.25)	(0, 1409)	(-17, 1409)
Assignment–Report	Mean	4.39	3.84	58.42	48.59
	Median	2.05	1.75	32.00	15.00
	Std. Dev.	5.65	4.84	67.31	98.48
	Range	(0, 33.19)	(0, 32.78)	(0, 343)	(-102, 1121)
Report–Technical Review	Mean	0.38	0.32	5.13	9.44
	Median	0.00	0.00	0.00	1.00
	Std. Dev.	0.92	0.83	11.93	21.29
	Range	(0, 7.06)	(0, 6.98)	(0, 73)	(-1, 319)
Tech Review–Admin Review	Mean	0.64	0.47	8.19	9.17
	Median	0.10	0.06	1.00	0.00
	Std. Dev.	1.32	0.96	15.54	30.17
	Range	(0, 9.19)	(0, 8.03)	(0, 95)	(0, 714)

Notes: Labor is defined as the number of staff reported for that year. Budget is defined as the annual DNA unit budget in $100,000 units. Turnaround time is reported in number of days.

Louisiana had a more stable baseline than the other study sites, making trends easier to identify among the non-outsourced cases (see figures 52–53,56). After the Lean Six Sigma process was first piloted, the laboratory was completing over six times more cases per month than before (a median of 15 cases per month before compared with 103 cases per month afterward) (figure 53). Moreover, the number of cases assigned[56] each month also quadrupled, suggesting that the new process not only improved the lab's ability to complete cases but also changed the capacity to start new work (figure 52). While there appears to be a small increasing trend before the pilot implementation (starting at the end of 2009), the shift post implementation is more dramatic.

Figures 54–55 show the annual *efficiency* throughput estimates.[57] The number of cases completed per month per analyst rose substantially in 2010, revealing that the Louisiana lab was able to truly increase its efficiency rather than merely producing more with additional staff resources. Although budget expenditures grew in 2010, the larger increase in monthly completed cases still meant that efficiency estimates more than doubled compared to previous years.

Figure 56 shows a reduction in turnaround time post-pilot compared to the previous three and a half years. However, it is unclear whether this change is due to the pilot or is part of the declining trend originating at the end of 2008. At the turn of this year, the lab instituted a few changes that might be responsible for the decreasing turnaround time. For instance, the lab began outsourcing samples in December 2008. While the figure only displays non-outsourced cases, it is possible that the practice of outsourcing allowed analysts to have a more manageable caseload, resulting in shorter turnaround times. At the beginning of 2009, the laboratory also began using electronic maintenance and quality control logs; however, it seems unlikely that this change would cause such a strong decreasing trend. Also at the beginning of the year, the lab started using a "tandem team" model where analysts batched samples together and divided the work by task (i.e., one analyst performed quantification while another performed amplification for all batched samples) instead of by case. Although this change was not widespread among staff and was not as strategically organized in the same manner as the LSS process, it is possible that this on its own or in combination with the other two events, began a trend of decreasing turnaround time. Other changes occurred after this point (see section 5.4.3, Regression Analyses, below for more details) that may have facilitated a continuing decline in turnaround time. The efficiency measures of turnaround time (figures 57–58) show greater efficiency in both 2009 and 2010 in terms of both labor and budget resources.

[56] Completed cases are not necessarily the same cases as those started each month. Started cases are matched to the month in which a case was assigned, while completed cases are assigned to the month in which the case was completed.

[57] Estimates are provided by year, because resource indicators were assess on an annual basis.

Louisiana did not electronically track many of the intermediary stages of DNA processing. However, the evaluators were able to document turnaround times for three of them: (1) assignment to report completion, (2) report completion to technical review completion, and (3) technical review completion to administrative review completion. When examining the assignment to report stage (figure 59), an additional decrease can be seen after the pilot, above and beyond the general declining trend starting in 2008. It seems possible that the grant activities did in fact reduce turnaround time for the bulk of DNA processing. However, these reductions are less visible when also including report review time. However, the time between report and technical review is so varied that is difficult to draw conclusions with such a modest follow-up period (figure 62). Although there may be an additional decrease for the administrative review after the pilot, it is small, and it is unclear whether the change is due to the pilot or to a more general decrease beginning earlier (figure 65). Stage-level efficiency measures all show similar patterns to those of the overall efficiency measures, with greater efficiency occurring in 2009 and 2010 compared to earlier years (see Figures 60–61, 63–64, and 66–67).

Finally, while other graphs depict the trends in non-outsourced cases, figure 68 shows the throughput of outsourced cases during the study period. A spike can be seen in summer 2010 when the laboratory conducted its focused technical review sessions to try to reduce their backlog of outsourced cases.

Figure 52. Monthly Number of Cases Started

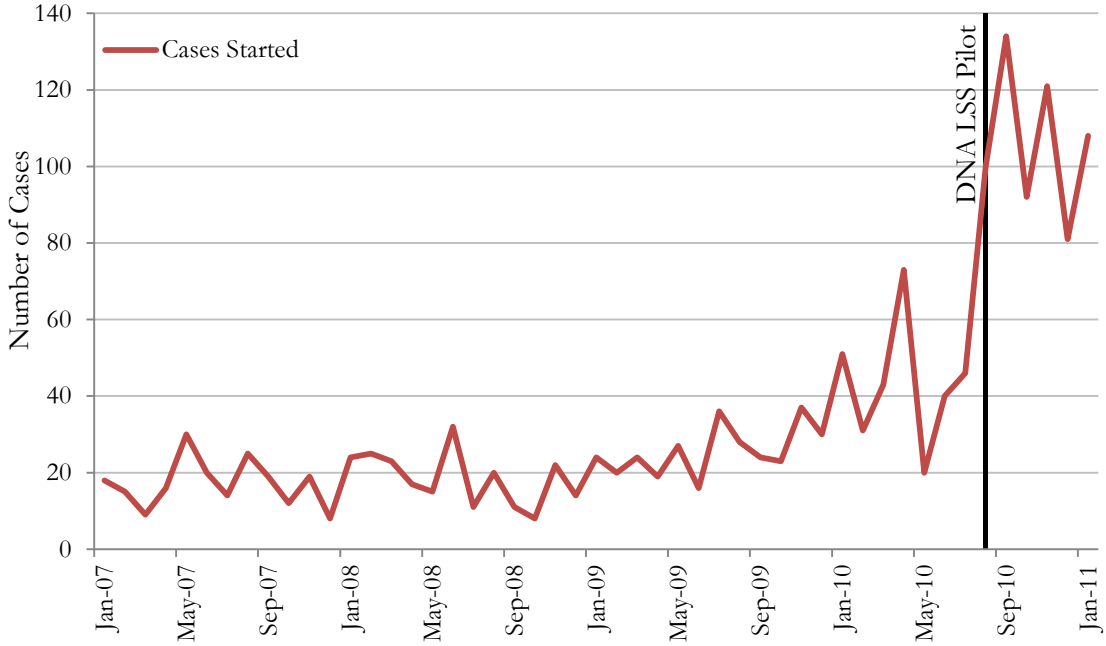


Figure 53. Monthly Throughput of Cases

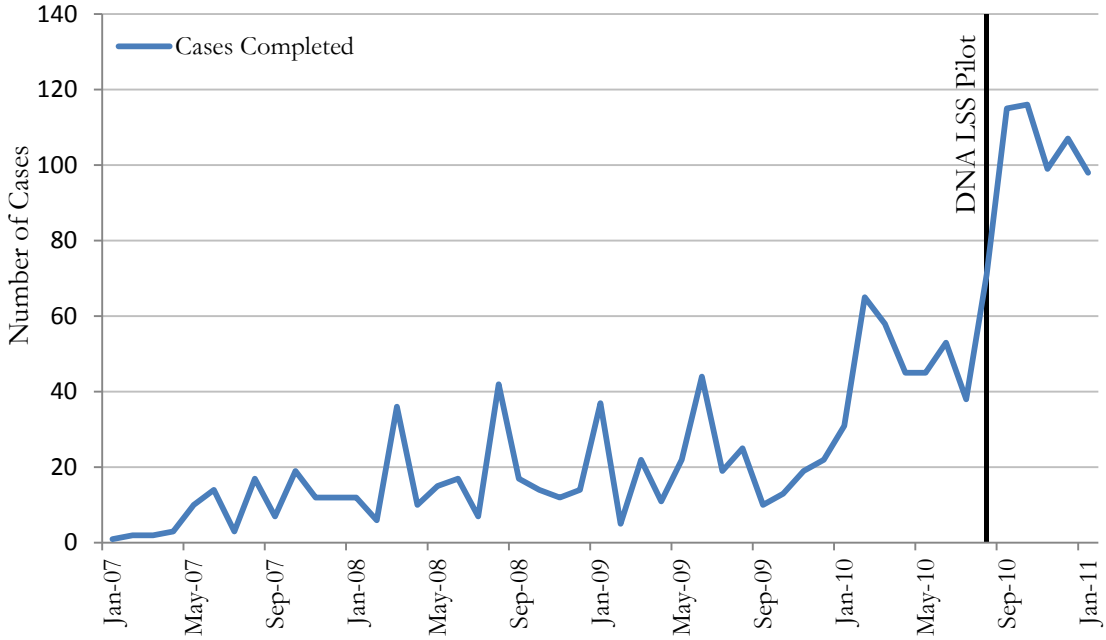


Figure 54. Efficiency Measure of Monthly Throughput by Labor

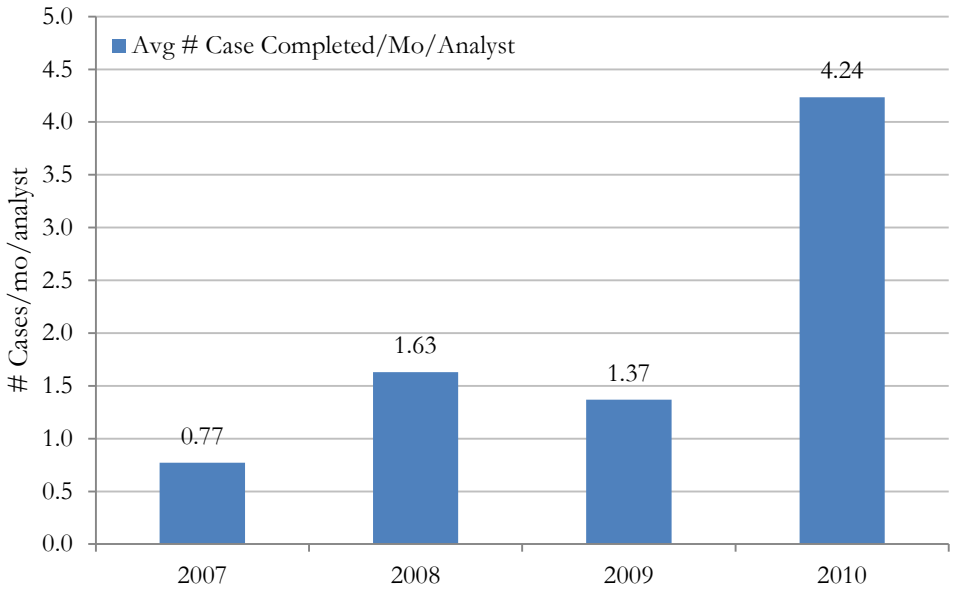
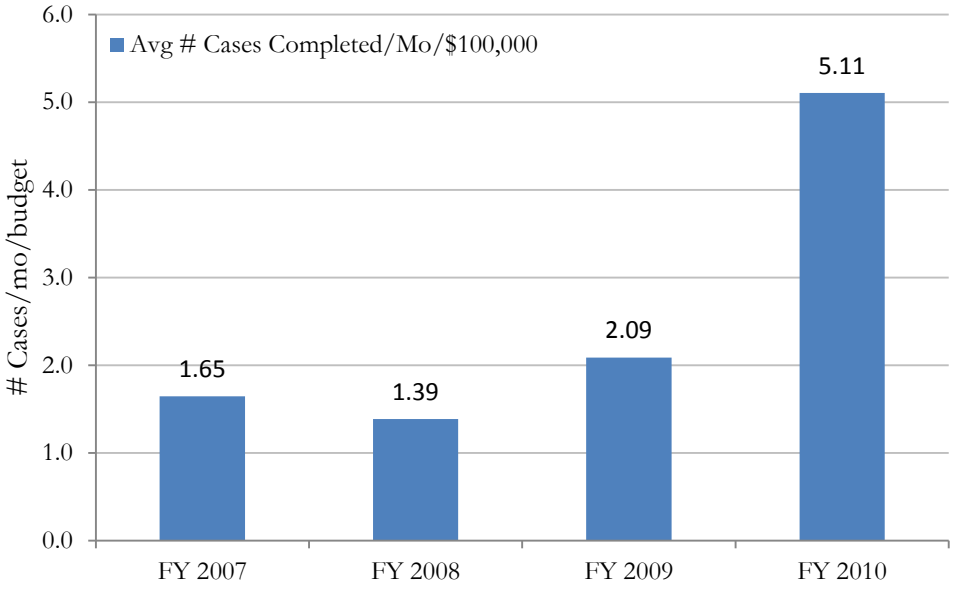


Figure 55. Efficiency Measure of Monthly Throughput by Budget Expenditures



Note: Because case processing data were not available for the last five months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 58 percent of the year.

Figure 56. Case Turnaround Time

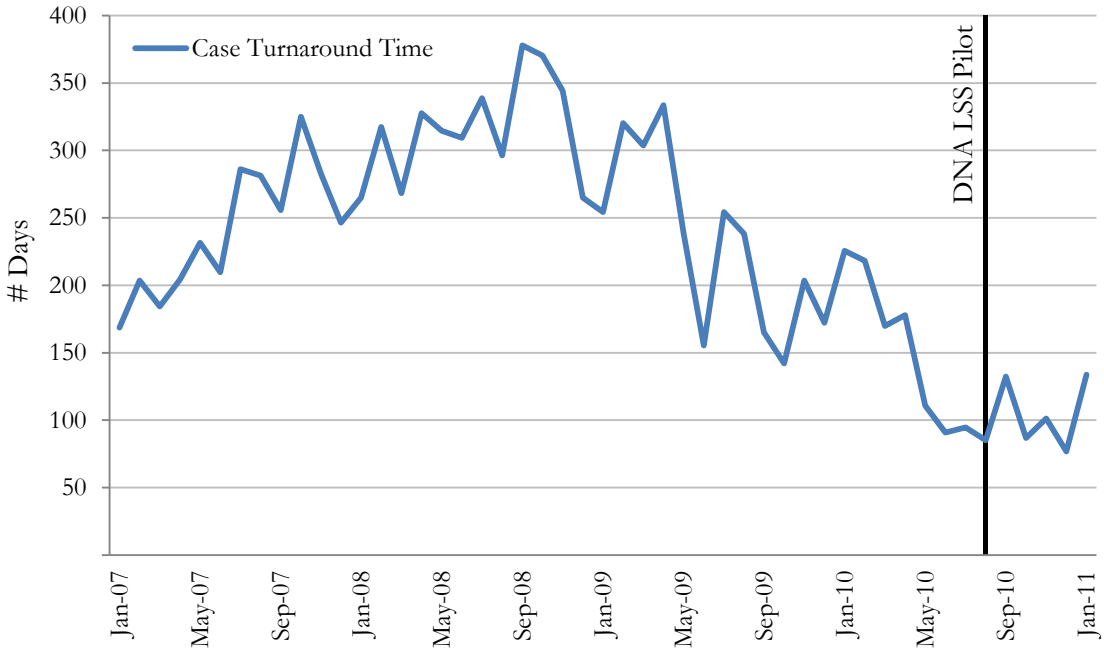


Figure 57. Efficiency Measure of Turnaround Time by Labor

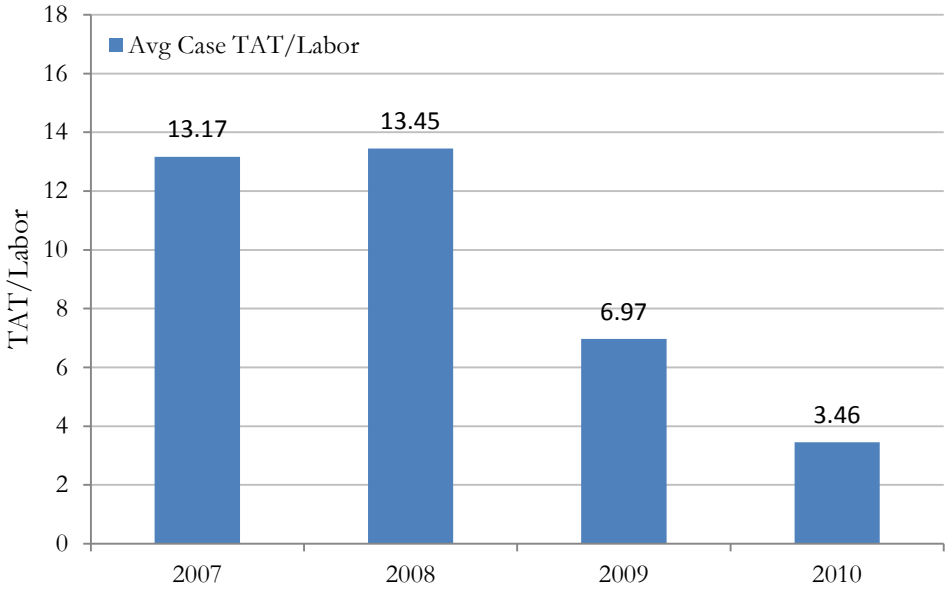
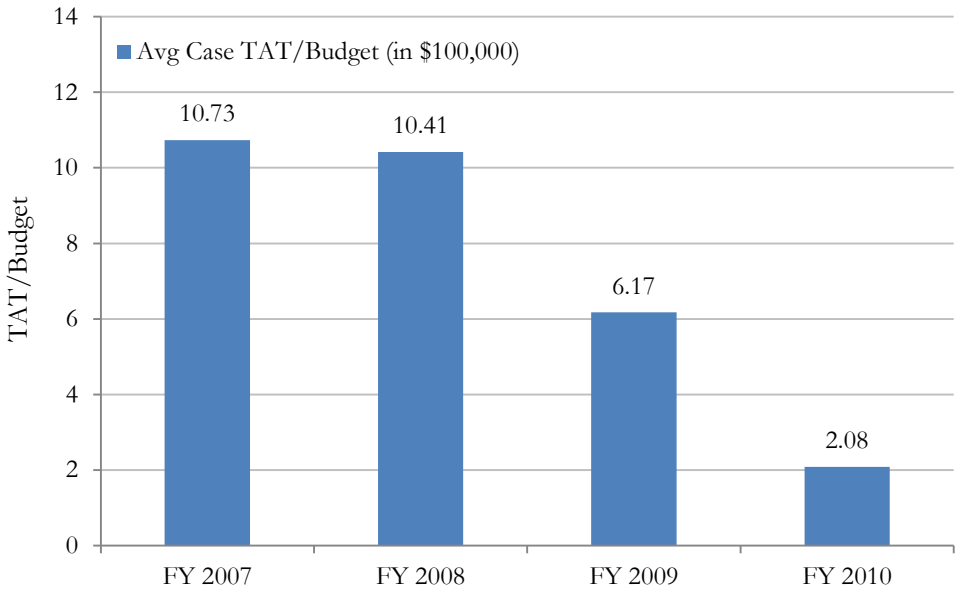


Figure 58. Efficiency Measure of Turnaround Time by Budget Expenditures



Note: Because case processing data were not available for the last five months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 58 percent of the year.

Figure 59. Stage-Level Turnaround Time: Assignment to Report

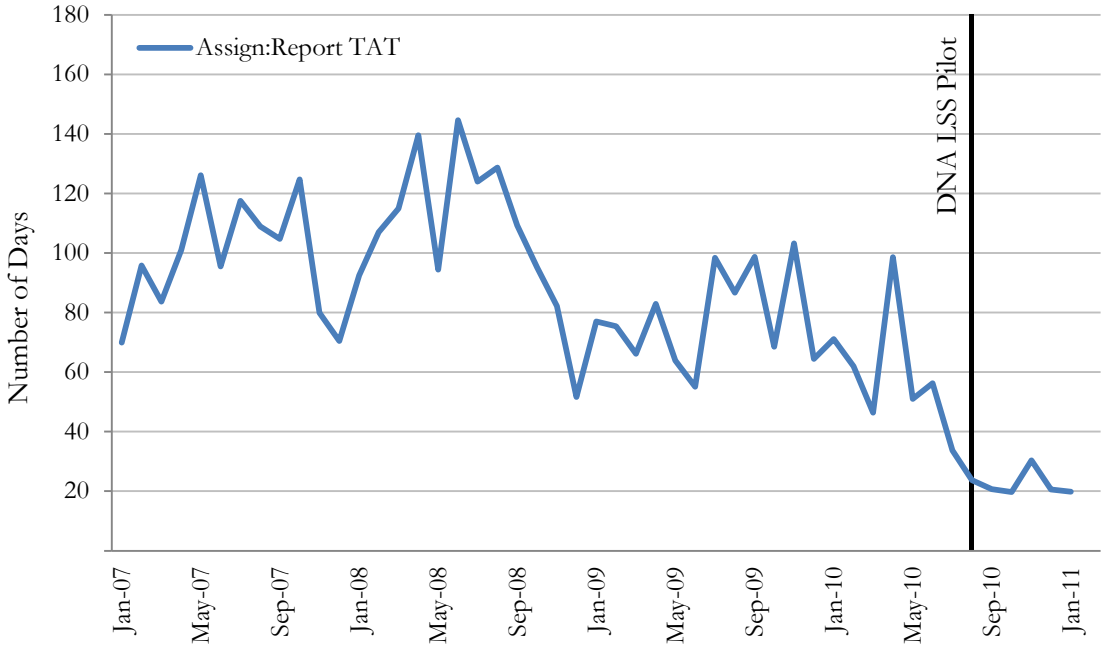


Figure 60. Stage-Level Efficiency Measure by Labor: Assignment to Report

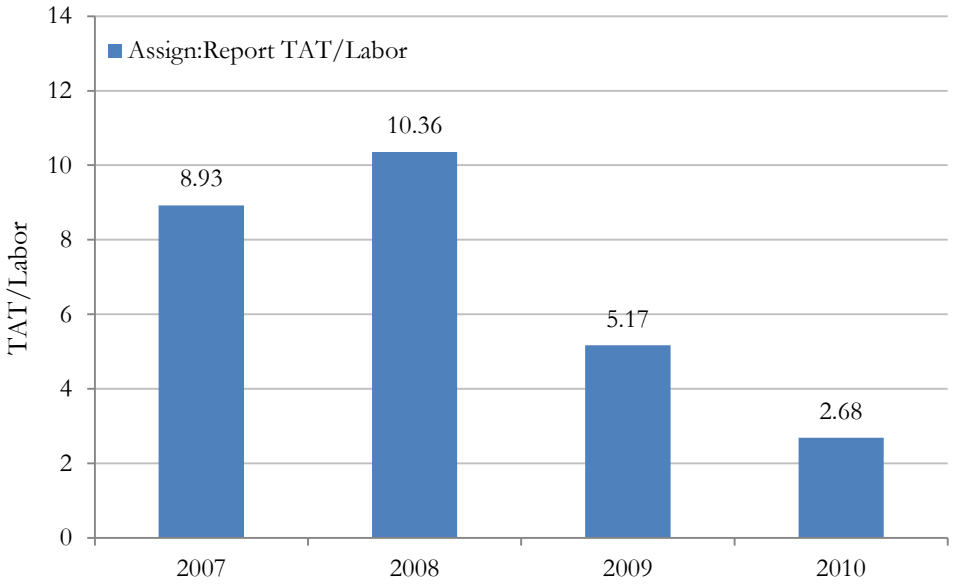
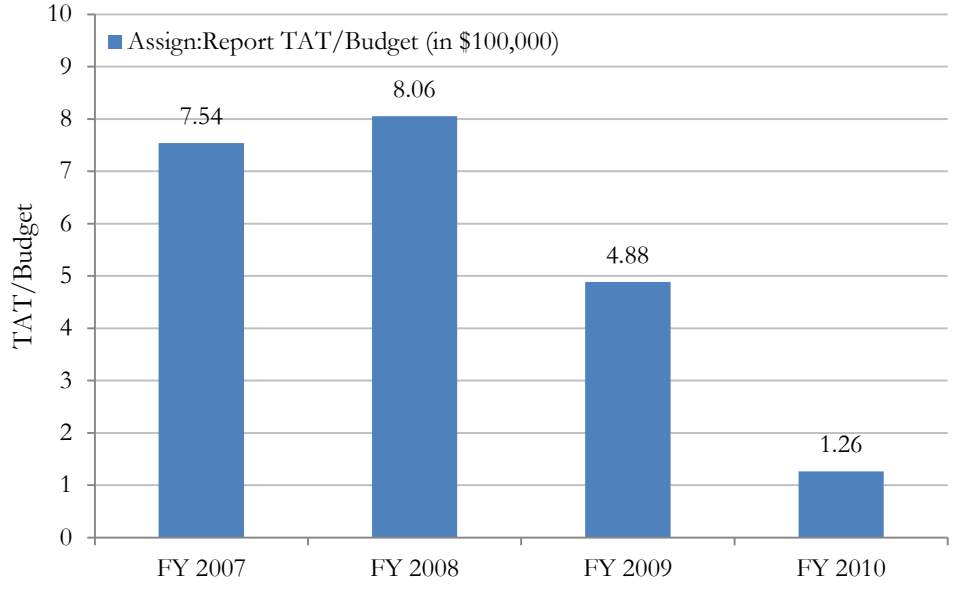


Figure 61. Stage-Level Efficiency Measure by Budget: Assignment to Report



Note: Because case processing data were not available for the last five months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 58 percent of the year.

Figure 62. Stage-Level Turnaround Time: Report to Technical Review

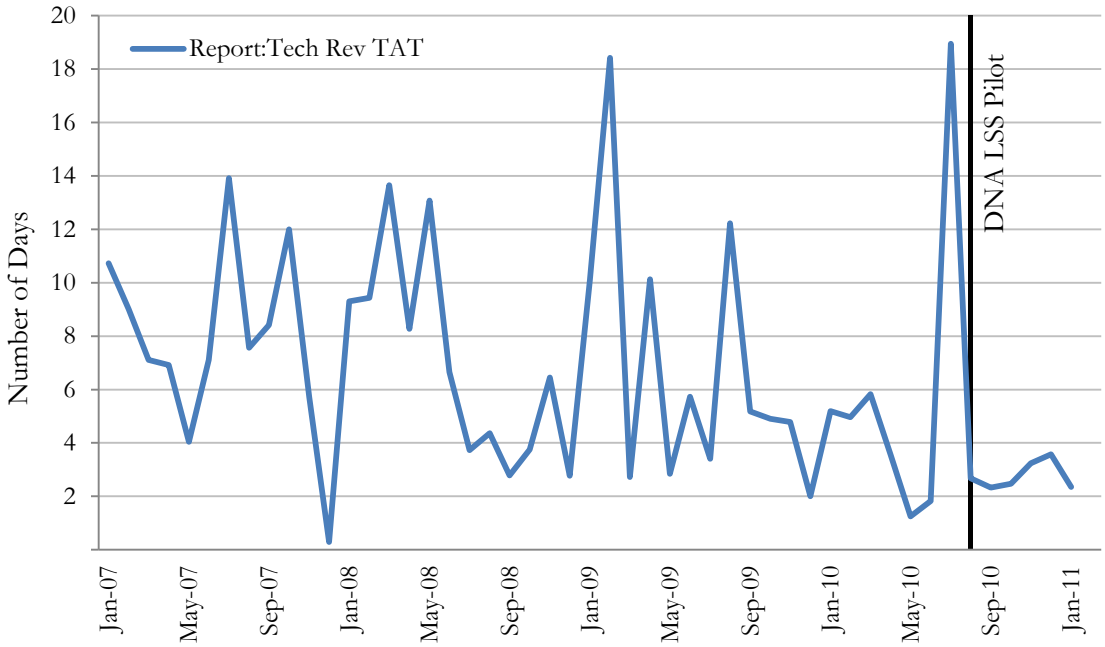


Figure 63. Stage-Level Efficiency Measure by Labor: Report to Technical Review

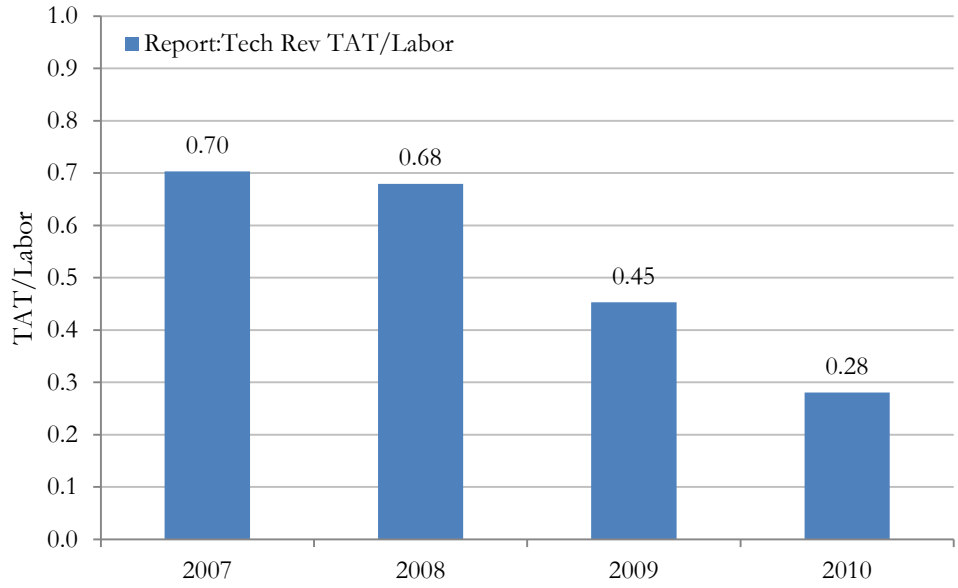
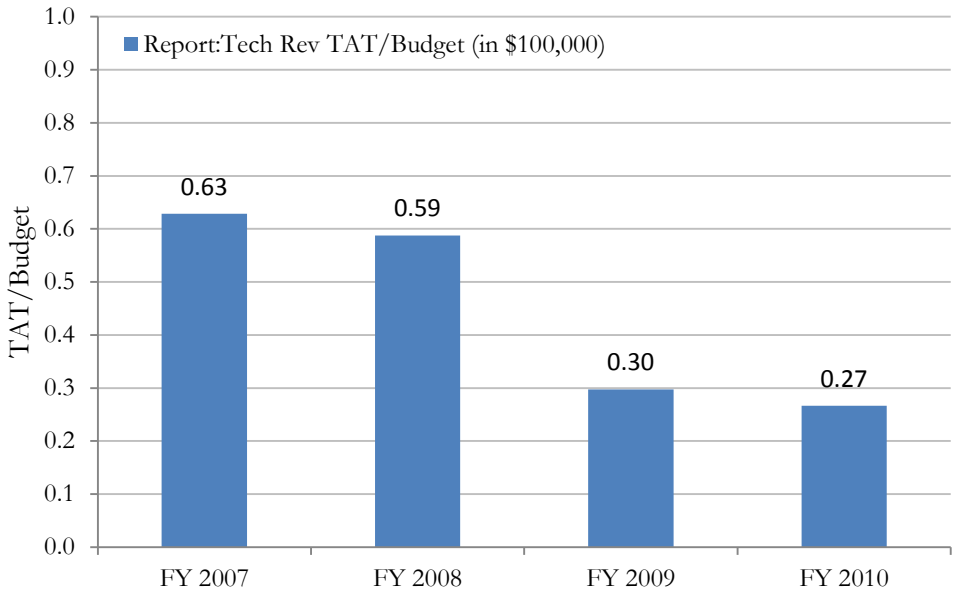


Figure 64. Stage-Level Efficiency Measure by Budget: Report to Technical Review



Note: Because case processing data were not available for the last five months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 58 percent of the year.

Figure 65. Stage-Level Turnaround Time: Technical Review to Administrative Review

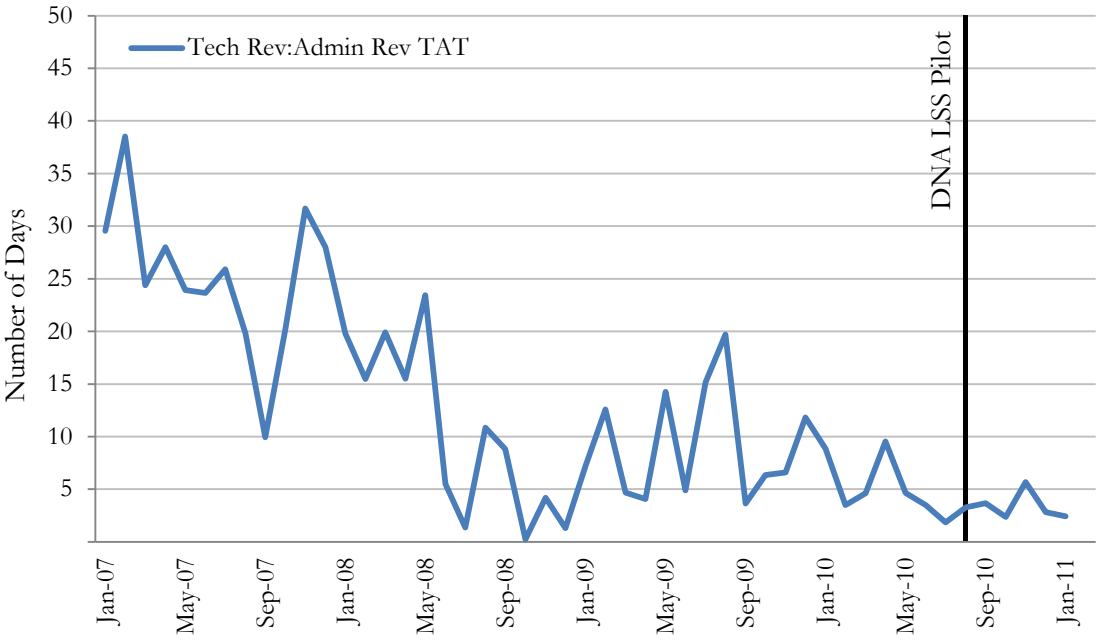


Figure 66. Stage-Level Efficiency Measure by Labor: Technical Review to Administrative Review

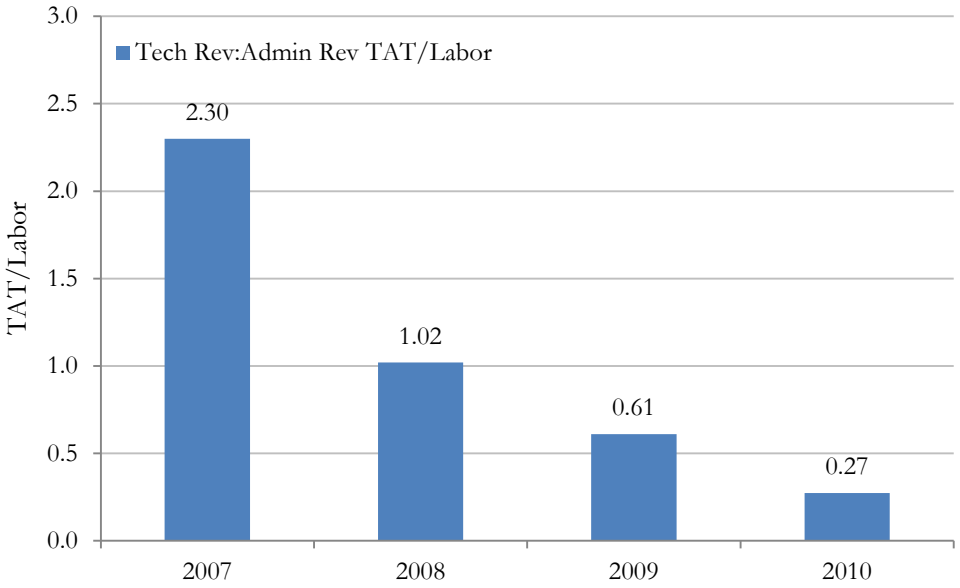
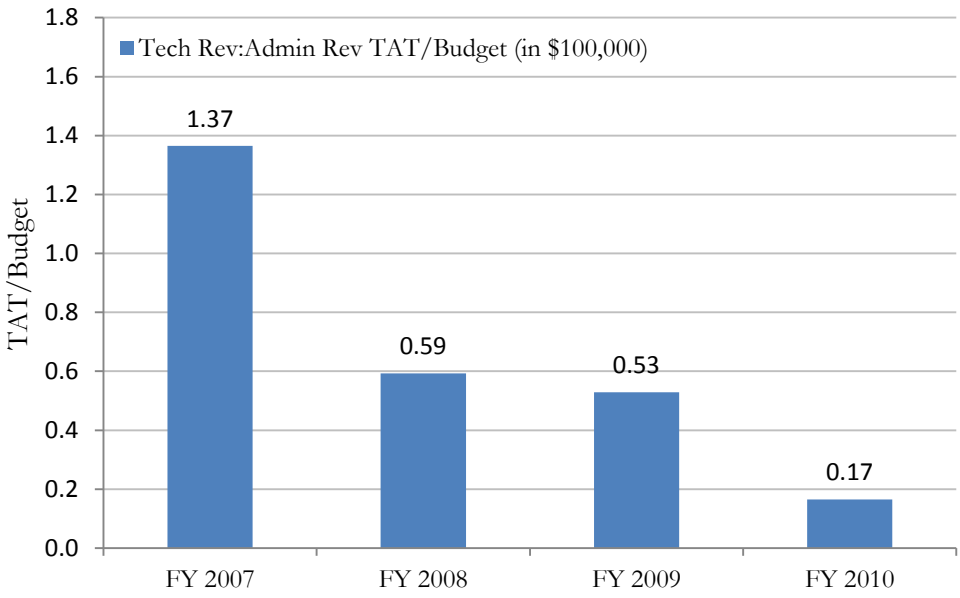
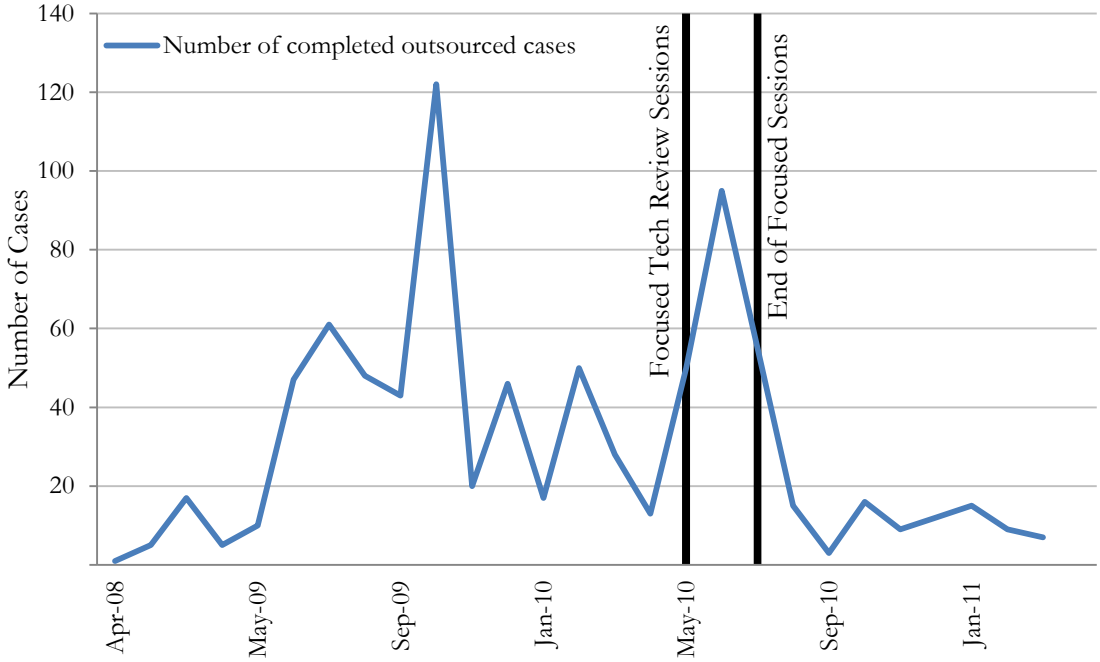


Figure 67. Stage-Level Efficiency Measure by Budget: Technical Review to Administrative Review



Note: Because case processing data were not available for the last five months of the 2010 fiscal year, the efficiency estimate may be underestimated since the budget is reported for an entire year, but case turnaround times are only provided for 58 percent of the year.

Figure 68. Throughput of Outsourced Cases



5.4.2 Pre/Post Comparisons

The research team conducted two types of tests (independent samples t-test and Mann-Whitney U test) to compare the throughput both before and after the first major implementation milestone: the piloting of the Lean Six Sigma process. From the pilot forward, the laboratory used the new Lean Six Sigma process to analyze all incoming DNA samples, although the process was modified throughout to tailor it better to the lab's work. Both the independent samples t-test and Mann-Whitney U test are used to compare differences between two groups. However, the Mann-Whitney U uses the median as the measure of central tendency and, therefore, does not require normality as the independent samples t-test does. Although t-tests are robust to violations of the normality assumption, the Mann-Whitney U test was also conducted since the data were skewed.

There was a significant increase in cases completed per month after the implementation of the new Lean Six Sigma process, as shown by results of both the t and U parameters tests (see table 9). This relationship remained significant after dividing the monthly throughput numbers by annual labor and budget expenditure to produce efficiency indices. Therefore, there is supportive evidence that the grant program had a positive impact on the Louisiana lab in terms of the number of cases completed per month, and this effect remains after taking into account labor and budget resources.

Table 9. Comparison Test Results for Louisiana

Throughput	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of LSS Pilot (Aug. 2010)	-11.33	<.001	0.00	<.001
Throughput/Labor	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of LSS Pilot (Aug. 2010)	-10.24	<.001	0.00	<.001
Throughput/Budget	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of LSS Pilot (Aug. 2010)	-8.61	<.001	1.00	<.001

5.4.3 Regression Analyses

In order to detect whether the program had an effect on turnaround time and its related efficiency measures after controlling for other case characteristics, the research team performed a series of negative binomial regression analyses (see table 10). Regressions were used to model overall and stage-level turnaround times, both as pure productivity measures (case turnaround time) and as efficiency measures (case turnaround time divided by the number of staff during the year the case began and case turnaround time divided by the year's budgetary expenditures in $100,000 units). The research team included a dummy variable that indicated whether the case occurred after the implementation of the LSS pilot on July 26, 2010. There were also four other event variables to control for the potentially confounding effects of other changes in the lab: (1) the start of outsourcing DNA cases in December 2008, use of electronic (as opposed to paper) maintenance and quality control logs in January 2009, and partial implementation of "tandem teams" in February 2009 where a team of analysts batched their samples together and divided processing responsibilities by task instead of by ownership of case (this occurred for 2–3 analysts); (2) the expansion of lab facility space and outsourcing of training in April 2009; (3) the acquisition of new computer, printer, and scanning equipment for barcode evidence tracking in July 2009; and (4) the production of electronic reports for laboratory clients and the creation of a Business Unit (to serve payroll and administrative support functions for entire lab instead of having DNA unit managers perform these tasks for their own unit) in January and February 2010, respectively. These potentially confounding events were coded in the same manner as the intervention variable.

Similar to the t-test and Mann-Whitney U findings, the analyses revealed a significant effect of the LSS process with shorter turnaround times once the new system was

implemented after accounting for other factors. According to the model, the Lean Six Sigma process reduced overall case turnaround by about 66 percent[58] (decreased turnaround time by 59 percent between assignment and report and 35 percent between report and technical review). The intervention had no apparent effect on the turnaround time between the completions of technical review and administrative review. These findings provide evidence that the NIJ grant program not only increased DNA case throughput, but also reduced the time spent processing a DNA case.

Other events at the lab also had significant relationships with turnaround time. In December 2008, the laboratory began outsourcing backlogged cases. In the two months after this change, the lab also switched to electronic maintenance and quality control logs and began using a tandem team model (one that was less structured and routinized than the assembly-line system of the pilot). Regression analyses found that this event was significantly related to turnaround time, with shorter overall turnaround times more likely after this point. These three events did not appear to influence the stage between report and technical review, but they reduced time between the technical and administrative reviews. As can be seen in figure 56, a long-term decline begins around this time and continues throughout the end of the study period.

Other events (including the lab facility expansion, outsourced training, new equipment, electronic versions of reports, and creation of the Business Unit) occurring after these first changes were also significant factors in predicting turnaround time. However, it is not clear whether these events contributed to the continued decrease. It is interesting to note that no outsourced cases are included in the analytic sample. This indicates that the decrease was caused by one of the other changes occurring at the lab (such as the tandem team model) or that outsourcing may have the ability to improve turnaround time for *non*-outsourced casework (possibly because analysts can focus more diligently on their smaller caseload). While every "confound" event variable was significant in the model predicting overall turnaround time, new computer and scanning equipment was not related to any of the three stages on their own. The lab expansion and outsourced training was only related to decreasing turnaround times for the assignment to report stage, and the institution of electronic reports and the Business Unit was related to decreasing turnaround times for the report to technical review stage only.

Other factors explored in the regression analyses included the number of items, type of evidence, related offense, and analyst experience. The presence of more evidence items per case was related to longer turnaround times for the overall case, assignment to report stages, and technical to administrative review stages. However, cases with more evidence items were associated with less time spent between the report and technical review. The type of evidence

[58] Negative binomial regressions are interpreted as the percent change in *y*, given a one-unit increase in *x*.

included in a case also was linked to turnaround time. Cases with known standard samples tended to have longer turnaround times at the overall and stage levels, with the exception of shorter turnaround times for the report to technical review stage. This increase in turnaround time for cases with samples generally assumed to have easier and faster processing may be due to the fact that known standards may be submitted at later dates, resulting in additional case time spent waiting on samples. Cases with intimate samples were associated with shorter turnaround times at all levels and stages (possibly due to the organized collection protocols leading to reduced screening time), while textile evidence was associated with longer turnaround times (possibly due to extra time spent "searching" for possible biological stains) for the overall case and the large stage before reviews. Cases with other types of swabs did not have a consistent relationship with turnaround time across the different stages and overall.

Property offenses were associated with shorter overall case turnaround time but longer time spent between the report and technical review. More experienced staff spent less time on a case, and these time savings were primarily found in the administrative review stage (greater experience was associated with more time spent during the technical review). This relationship with analyst experience is puzzling since the technical and administrative reviews are conducted by other staff members. However, it is possible that experienced analysts have fewer clerical and administrative errors, which might explain the shorter time spent during this review. The stage where the analyst would have most responsibility (assignment to report) does not reveal a significant relationship.

The results for efficiency measures for routing turnaround time were similar to those for the productivity measures. Exceptions were that (1) property offenses were no longer significantly related to turnaround time, (2) new computer and scanning equipment was associated with reduced turnaround time between the assignment and report stages once budget expenditures were accounted for, and (3) the intervention became a significant predictor of turnaround time between the technical and administrative reviews after taking into account budget expenditures. In addition, there were several changes in the two review stage models regarding item number and type, as well as analyst experience once laboratory resources were taken into account.

Table 10. Regression Results for Louisiana

Overall TAT Regression	**Overall Case TAT**		**Overall Case TAT/Labor**		**Overall Case TAT/Budget**	
	b coefficient	*p-value*	*b coefficient*	*p-value*	*b coefficient*	*p-value*
Intervention: LSS Pilot	-0.66	<.01	-0.65	<.01	-0.86	<.01
Confound: Outsourcing of DNA cases, electronic logs, & tandem teams	-0.61	<.01	-0.95	<.01	-0.36	<.01
Confound: Lab facility expansion & outsourcing of training	-0.50	<.01	-0.50	<.01	-0.49	<.01
Confound: New computer and scanning equipment	0.29	<.01	0.29	<.01	-0.10	0.17
Confound: Electronic reports & Business Unit	-0.17	<.01	-0.25	<.01	-0.17	<.01
Number of items in case	0.02	<.01	0.01	<.01	0.02	<.01
Item Type: Standards	0.37	<.01	0.38	<.01	0.38	<.01
Item Type: Intimate Samples	-0.45	<.01	-0.38	<.01	-0.41	<.01
Item Type: Textiles	0.20	<.01	0.20	<.01	0.21	<.01
Item Type: Other Swabs	-0.03	0.54	0.01	0.79	0.00	0.94
Violent Offense	0.00	0.96	0.00	0.97	-0.02	0.80
Property Offense	-0.15	0.02	-0.11	0.08	-0.12	0.07
Analyst Experience	-0.03	<.01	-0.03	<.01	-0.03	<.01
Stage-Level Regression	**Assign–Report TAT**		**Report–Tech Rev TAT**		**Tech Rev–Admin Rev TAT**	
	b coefficient	*p-value*	*b coefficient*	*p-value*	*b coefficient*	*p-value*
Intervention: LSS Pilot	-0.59	<.01	-0.35	<.01	-0.02	0.91
Confound: Outsourcing of DNA cases, electronic logs, & tandem teams	-0.67	<.01	-0.06	0.80	-1.27	<.01

Stage-Level Regression	Assign–Report TAT		Report–Tech Rev TAT		Tech Rev–Admin Rev TAT	
Confound: Lab facility expansion & outsourcing of training	-0.47	0.01	-0.10	0.71	-0.30	0.33
Confound: New computer and scanning equipment	-0.11	0.36	0.22	0.25	-0.17	0.43
Confound: Electronic reports & Business Unit	0.06	0.40	-0.57	<.01	-0.15	0.31
Number of items in case	0.04	<.01	-0.02	<.01	0.02	<.01
Item Type: Standards	0.61	<.01	-0.46	<.01	0.33	0.02
Item Type: Intimate Samples	-0.50	<.01	-0.25	0.03	-0.36	<.01
Item Type: Textiles	0.21	<.01	-0.12	0.34	-0.26	0.06
Item Type: Other Swabs	0.12	0.04	-0.41	<.01	0.35	<.01
Violent Offense	-0.20	0.03	0.43	<.01	-0.19	0.24
Property Offense	-0.27	<.01	0.57	<.01	-0.17	0.36
Analyst Experience	0.02	0.06	0.04	0.04	-0.11	<.01
Stage-Level Regression	**Assign–Report TAT/Labor**		**Report–Tech Rev TAT/Labor**		**Tech Rev–Admin Rev TAT/Labor**	
	b coefficient	*p-value*	*b coefficient*	*p-value*	*b coefficient*	*p-value*
Intervention: LSS Pilot	-0.57	<.01	-0.36	<.01	-0.15	0.31
Confound: Outsourcing of DNA cases, electronic logs, & tandem teams	-1.02	<.01	-0.36	0.06	-1.50	<.01
Confound: Lab facility expansion & outsourcing of training	-0.48	<.01	-0.19	0.39	-0.33	0.22
Confound: New computer and scanning equipment	-0.08	0.44	0.29	0.07	-0.10	0.61
Confound: Electronic reports & Business Unit	-0.02	0.79	-0.62	<.01	-0.24	0.11
Number of items in case	0.03	<.01	-0.02	<.01	-0.002	0.61

Stage-Level Regression	Assign–Report TAT/Labor		Report–Tech Rev TAT/Labor		Tech Rev–Admin Rev TAT/Labor	
Item Type: Standards	0.61	<.01	-0.46	<.01	0.06	0.60
Item Type: Intimate Samples	-0.36	<.01	-0.14	0.16	-0.21	0.06
Item Type: Textiles	0.21	<.01	-0.11	0.29	-0.21	0.10
Item Type: Other Swabs	0.12	0.02	-0.42	<.01	0.16	0.08
Violent Offense	-0.17	0.02	0.41	<.01	0.03	0.84
Property Offense	-0.25	<.01	0.59	<.01	-0.06	0.69
Analyst Experience	0.02	0.06	0.02	0.30	-0.11	<.01

Stage-Level Regression	Assign–Report TAT/Budget		Report–Tech Rev TAT/Budget		Tech Rev–Admin Rev TAT/Budget	
	b coefficient	*p-value*	*b coefficient*	*p-value*	*b coefficient*	*p-value*
Intervention: LSS Pilot	-0.78	<.01	-0.56	<.01	-0.37	0.01
Confound: Outsourcing of DNA cases, electronic logs, & tandem teams	-0.43	<.01	0.36	0.04	-0.63	<.01
Confound: Lab facility expansion & outsourcing of training	-0.47	<.01	-0.16	0.40	-0.32	0.16
Confound: New computer and scanning equipment	-0.47	<.01	-0.11	0.44	-0.32	0.16
Confound: Electronic reports & Business Unit	0.06	0.37	-0.54	<.01	-0.48	0.01
Number of items in case	0.03	<.01	-0.02	<.01	-0.16	0.27
Item Type: Standards	0.60	<.01	-0.45	<.01	0.002	0.60
Item Type: Intimate Samples	-0.38	<.01	-0.07	0.48	-0.30	0.01
Item Type: Textiles	0.22	<.01	-0.09	0.39	-0.24	0.08
Item Type: Other Swabs	0.12	0.02	-0.37	<.01	0.24	0.02
Violent Offense	-0.18	0.02	0.42	<.01	-0.10	0.47
Property Offense	-0.25	<.01	0.65	<.01	-0.22	0.16
Analyst Experience	0.02	0.10	0.02	0.19	-0.10	<.01

5.4.4 Conclusions

The Louisiana State Police Crime Lab intended to improve the efficiency of forensic DNA evidence processing through hiring consultants to recommend changes to general casework and purchasing processes at the lab, acquiring a DLIMS module, and obtaining and/or validating various other technologies and chemical procedures. The two consulting groups both used the Lean Six Sigma methodology to improve routines at the lab. Lean Six Sigma is a data-driven approach that emphasizes removing unnecessary and wasteful activities, instituting an assembly-line approach, increasing accountability, and using performance measurement as a tool for assessment and improvement. Through a series of activities to assess and measure current problems or areas for improvement, solutions were developed to address targeted issues. The lab attempted to purchase a DLIMS module for the lab but was unable to identify a suitable option within the given budget. Instead , funds were redirected to other lab improvements, such as the Lean Six Sigma consulting contract for purchasing services and a transition to electronic data and document management, as well as contributing to the lab's new barcode tracking. Other instituted technologies and procedures included robotics, new extraction chemistries, and bone extraction equipment, among others. Although all of these were implemented by the end of the study period, only the original Lean Six Sigma for DNA casework occurred before data collection. Therefore, the analyses focused on the effects of this particular intervention.

Louisiana had significant delays in the initiation of its project due to competing obligations from other grants, but the lab quickly made up for lost time. The lab also reported other implementation challenges, including purchasing delays, inability to identify a suitable DNA LIMS module, and the difficulty of incorporating new grant activities (such as the Lean Six Sigma exercises) into staff time while continuing to maintain normal casework operations.

Louisiana had a more stable baseline than the other study sites, making trends easier to identify. The laboratory was completing over six times more cases per month compared to before the institution of Lean Six Sigma. The new process not only improved the lab's ability to complete cases but also changed the capacity to start new work. Pre/post comparison tests of throughput confirm graphical findings. There was a significant increase in cases completed per month after the implementation of the new LSS process. Findings were similar for both labor and budget efficiency indices.

Effects of the grant program on turnaround time were less clear. Trend data show a continued decline after LSS was implemented (decreasing turnaround time by about 66 percent); however, it is unclear whether this is due to the grant or to a previously existing declining trend beginning in 2008. When broken into stages, a more distinct break can be seen after the LSS pilot for the assignment to report stage. However, no clear pattern can be

determined from the report and review stages. This may account for why the overall turnaround time does not show as clear a decrease after the LSS implementation.

Efficiency indices also revealed more efficient DNA processing in 2010 for throughput and turnaround time in terms of both staff and financial resources. Meanwhile, other lab events and case characteristics also were related to turnaround time. A series of nongrant-related lab changes may have also decreased turnaround time, although it is difficult to determine if each event contributed uniquely to the decline that began in 2008, or if these events happened to align with a preexisting trend. Other factors were related to turnaround time as well. For instance, a smaller number of items, property offenses, and analyst experience were related to shorter overall turnaround time. Type of evidence had a varying relationship with longer turnaround time associated with textiles (possibly due to searching) and known standards (possibly due to the greater likelihood of multiple submissions and related delays in evidence receipt), while shorter turnaround times were related to intimate sample types (possibly due to the more standardized nature of sexual assault kits).

Overall, there is supportive evidence that the grant program had a positive and strong impact on the Louisiana lab in terms of the number of cases completed per month and turnaround time. This effect also remains after taking into account labor and budget resources. Due to delayed implementation of the lab's other grant components, the current study could not evaluate the effects of the lab's additional interventions, but it is possible that these changes have created additional gains in productivity and efficiency.

6. CASE STUDY: SAN FRANCISCO POLICE DEPARTMENT CRIMINALISTICS LABORATORY

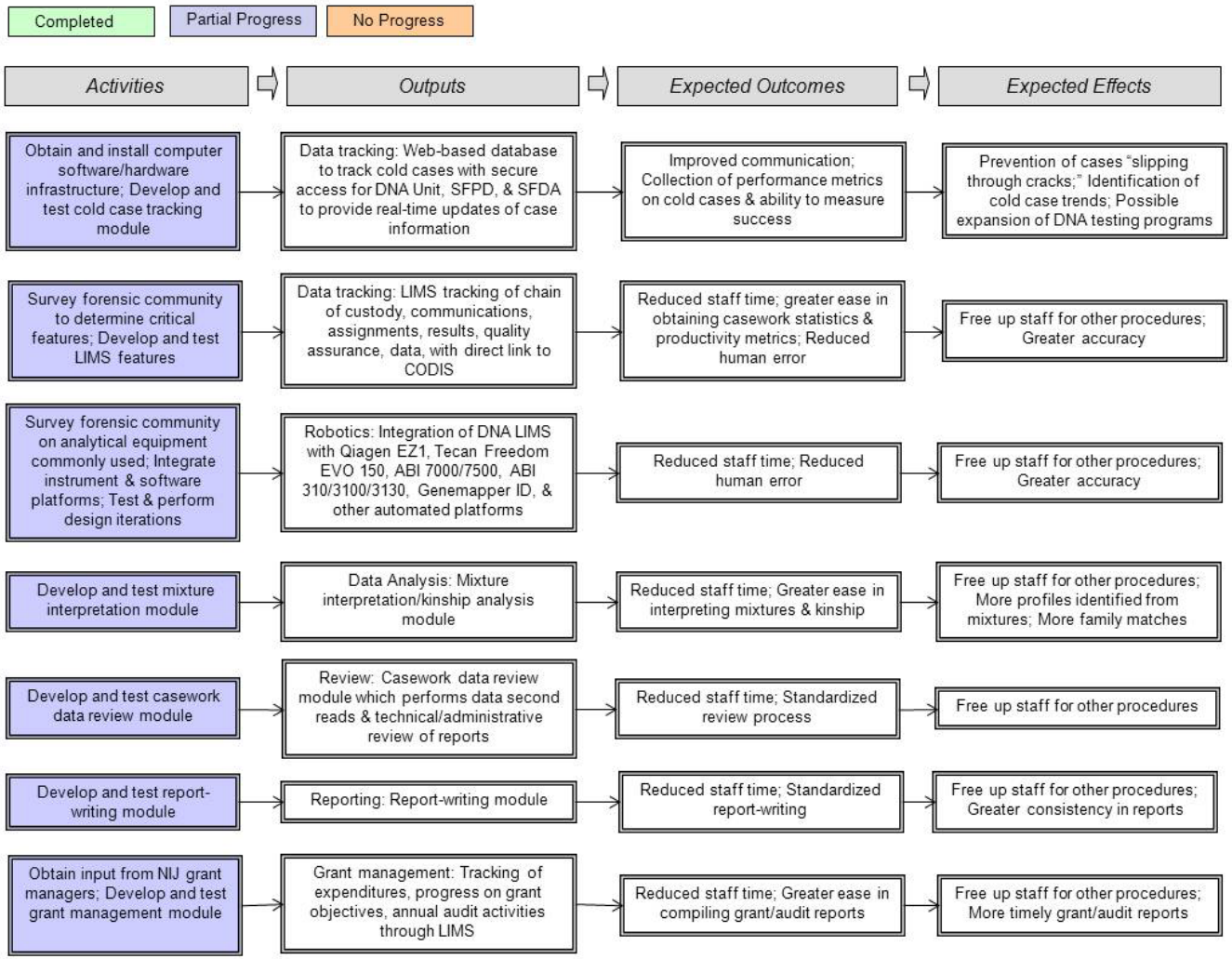
6.1 Overview of the Laboratory

San Francisco Police Department Criminalistics Laboratory (hereafter, San Francisco) is an accredited public crime laboratory that accepts crime scene and known reference DNA samples from both the city and county of San Francisco. For the NIJ grant, San Francisco proposed to develop a comprehensive case management system, with modules for cold hits, mixture interpretation, administrative tasks, data review, report-writing, quality assurance, and grant management, that could be used by its laboratory and other DNA laboratories.

6.2 Description of Grant Goals

NIJ awarded San Francisco a $1,024,467 grant to fund personnel costs, travel expenses, vendor services to customize the database, and the purchase of hardware and software needed to migrate current LIMS data onto a new server. San Francisco's 25 percent nonfederal match supported personnel costs and the purchase of the base LIMS product (to be customized further by vendor). San Francisco described eight main modules of the proposed Forensic Case Management System, or FMS (see figure 69). The goals of each of these modules, the activities involved in producing these modules, and the expected outcomes and impacts are described below. It was expected that achieving these goals would result in greater efficiency in DNA processing.

Figure 69. Logic Model for San Francisco



The primary goal of the grant was to create a web-based cold-hit tracking module that provides the DNA unit, Police Department, and District Attorney's Office with secure access to information in real time about cases involving "cold hits" or DNA matches made to unknown individuals through DNA databases. The module would track various case milestones after the hit, such as locating the suspect, arrest, arraignment, and case outcome. The overall success of DNA databases such as CODIS relies on effective police and prosecution follow-up once a hit is made. This cold-hit module would improve communication between forensics, police, and prosecutors. Further, it provided the ability to track cold-hit outcomes and, therefore, collect performance metrics on the DNA unit's success with cold cases. This could lead to the identification of cold-hit trends, an expansion of DNA testing programs due to better performance measurement, and hopefully would prevent cases from "slipping through the cracks" after a hit is made.

In addition to the cold-hits module, the FMS would include administrative casework features typical of many LIMS databases, including tracking of chain of custody, communications, assignments, and storage of results and profiles. The FMS would also have a direct link to CODIS for easy uploading of DNA profiles. A quality assurance module would track quality control measures, and a grant management module would track expenditures, progress on grant objectives, and annual audit activities. It was thought that an organized LIMS database would reduce staff time spent on administrative tasks, decrease human error, facilitate the tracking of samples and quality assurance issues, and assist in compiling grant or audit reports and obtaining casework statistics and productivity metrics.

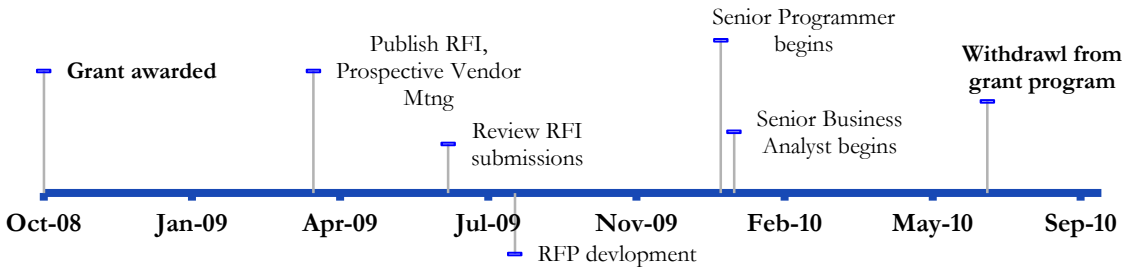
The FMS would also include a module to assist in data analysis, providing kinship analysis and interpretation of mixtures. Such a module would lead to easier interpretation of mixtures and kinship relationships and a reduction in staff time spent on interpretation. Once the data were analyzed and interpreted, a report-writing module would help to reduce staff time spent on writing reports and standardize the formats of reports within the lab. A casework data review module documents secondary reviews of data and technical/administrative review of reports. This would also reduce staff time spent on review activities and standardize the review process.

The FMS would be integrated with common lab equipment and software to facilitate the automated transfer of data between systems. Automated data transfer was expected to lead to reduced staff time spent on manually moving data and a reduction in human error. San Francisco anticipated the system to be completed within five years.

6.3 Implementation Findings

6.3.1 Implementation Description

Figure 70. San Francisco Implementation Timeline



In order to create this comprehensive Forensic Case Management System, which would include a grant management module to assist lab managers with the grant organization, San Francisco administered a national survey (N = 50) and state survey (N = 12) to learn the forensic community's needs regarding LIMS databases. They also inquired about software (e.g., *Genemapper ID*) and lab equipment (e.g., robots, capillary electrophoresis instruments) in order to identify common lab equipment that would need to be integrated with the system for use at other labs. The next step was for San Francisco to select a vendor through a competitive bid process. In order to help with the bid process, San Francisco hired a senior business analyst to help develop the solicitation announcement. The lab also hired a senior programmer analyst and a network administrator to interact with the vendor, prepare necessary infrastructure, and test the functionality of the new system.

San Francisco then needed to obtain and install the necessary hardware and software for the system, and develop and test the various modules. For the grant management module, San Francisco also wanted to obtain input from NIJ on important features to include. San Francisco expected there to be a substantial process of developing, testing, and design iterations, including a period of beta testing within other public laboratories.

6.3.2 Implementation Challenges

San Francisco experienced a significant delay in grant activities due to an extremely long vendor selection process. The Request for Information was not published until April 2009. They held a meeting for about 12 prospective vendors during the same month. Work continued with the drafting and repeated editing of the Request for Proposals for almost 12

months. While the RFP was being developed, San Francisco worked to hire the other staff needed for their project, including a senior programmer analyst, network administrator, and senior business analyst.

The project manager also reported challenges with hiring of internal staff. While the researchers could not independently confirm, the local point of contact suggested that although funds were available through the grant to support the hiring of some staff members important to the project's development, the project manager said it was difficult to convince the police department to hire new staff when they were in the midst of budget problems and were cutting other existing positions.

Ultimately, the Request for Proposals was never published. San Francisco concluded their participation in grant activities June 2010, after withdrawing from the grant program. All proposed grant activities ceased and all remaining grant funds were returned to NIJ. This occurred due to events external to the grant program that included both the loss of key personnel and the demands of participating in outside audits of the controlled substances and DNA unit sections.

6.4 Outcome Findings

Since there was no DNA-processing data collected from this site, there are no outcome results to report.

7. CASE STUDY: UNIVERSITY OF NORTH TEXAS HEALTH SCIENCES CENTER AT FORT WORTH

7.1 Overview of the Laboratory

The University of North Texas Health Sciences Center at Fort Worth (hereafter, UNT) houses the Center for Human Identification, an accredited public laboratory whose primary purpose is to identify human remains. The laboratory has a Forensic Casework division,[59] a Research and Development division (Field Testing division),[60] and a Paternity Testing division. The casework division has three units, one that performs DNA analysis on unidentified human remains, another which performs DNA analysis for county forensic cases, and the family reference sample team, which analyzes blood and buccal swab samples from family members of missing persons. Human remains are identified by comparing family DNA profiles to the DNA profiles obtained from human remains in the CODIS. The laboratory also partners with the University of North Texas Forensic Genetics Program to train graduate students.

The lab will accept samples from all states and also does some international DNA databasing work.[61] It is one of the few laboratories in the country that routinely processes human remains for mitochondrial DNA. Mitochondrial DNA (mtDNA) is inherited through the maternal bloodline, so it is used to make associations between family members (e.g., a mother, son, and daughter will all have the same mtDNA). Notable for the UNT case study, mitochondrial DNA is analyzed by sequencing the nucleotide bases of a particular region, rather than generating a profile of selected loci as in STR analysis. Mitochondrial DNA sequencing is more time consuming than traditional nuclear DNA STR processing (two weeks or more for mtDNA compared to about one week for STR). UNT is the second-largest forensic lab in the United States that analyzes mitochondrial DNA and supplies 45 percent of the entries into the CODIS Missing Persons index.

UNT proposed to implement a series of new approaches to analyze mitochondrial DNA family reference samples, including changes related to chemistry, robotics, expert filtering software, and data tracking. UNT's research division developed and validated these techniques for eventual implementation by the Family Reference Sample Team of the casework division.

[59] Hereafter, "casework division."

[60] Hereafter, "research division."

[61] UNT will process known standard samples that were collected in order to be uploaded to a DNA database.

7.2 Description of Grant Goals

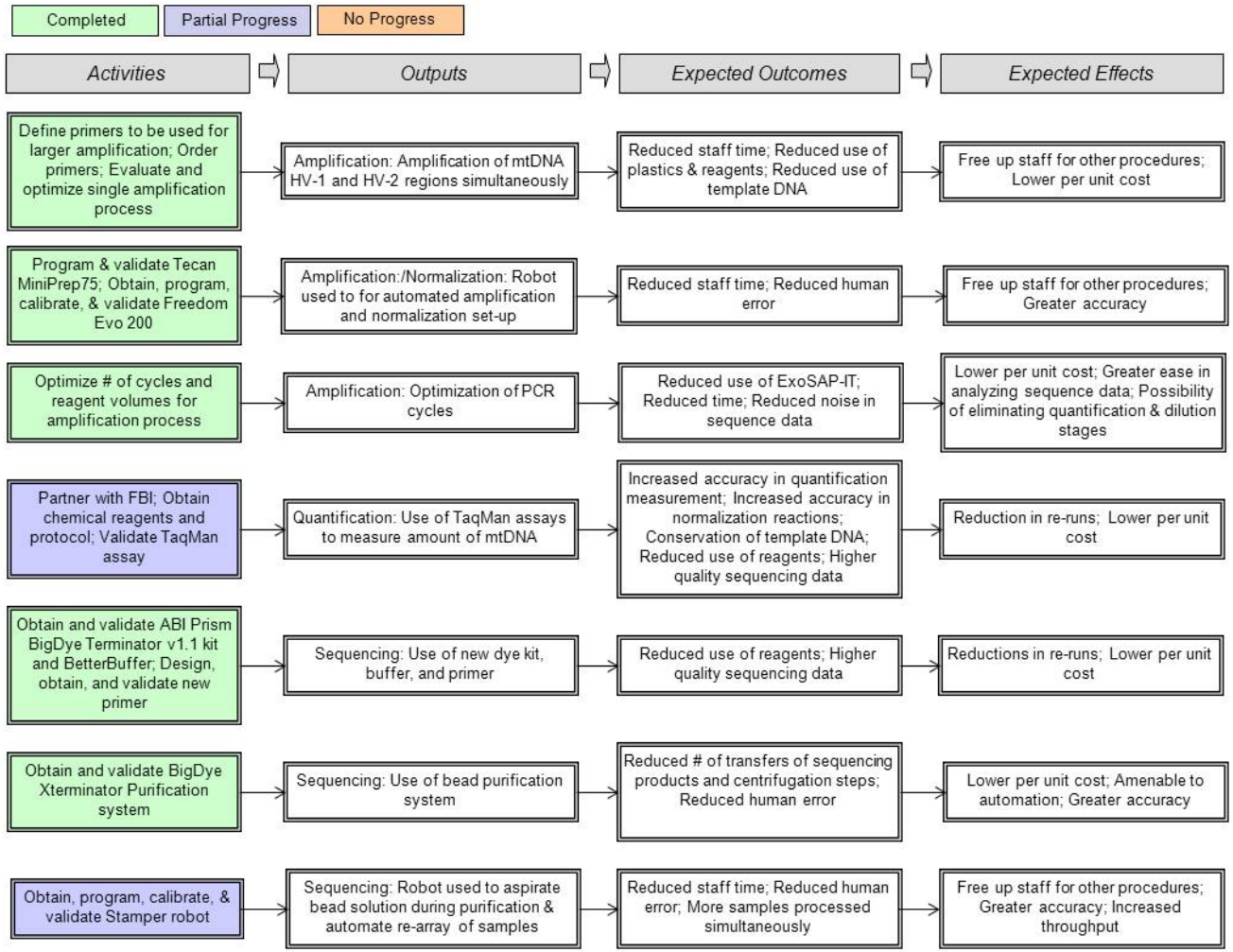
NIJ awarded UNT a $601,632 grant to fund robotics equipment, reagent and consumable supplies, software development and consultation, and indirect costs. UNT's 25 percent nonfederal match supported personnel costs and conference attendance related to the grant activities. UNT described 13 main goals of the grant to improve the efficiency of processing mitochondrial DNA family reference samples within its lab. These goals, the activities involved in achieving these goals, and the expected outcomes and impacts are described below (see figure 71). UNT predicted the implemented changes would increase the number of mitochondrial DNA samples processed by 35 percent.

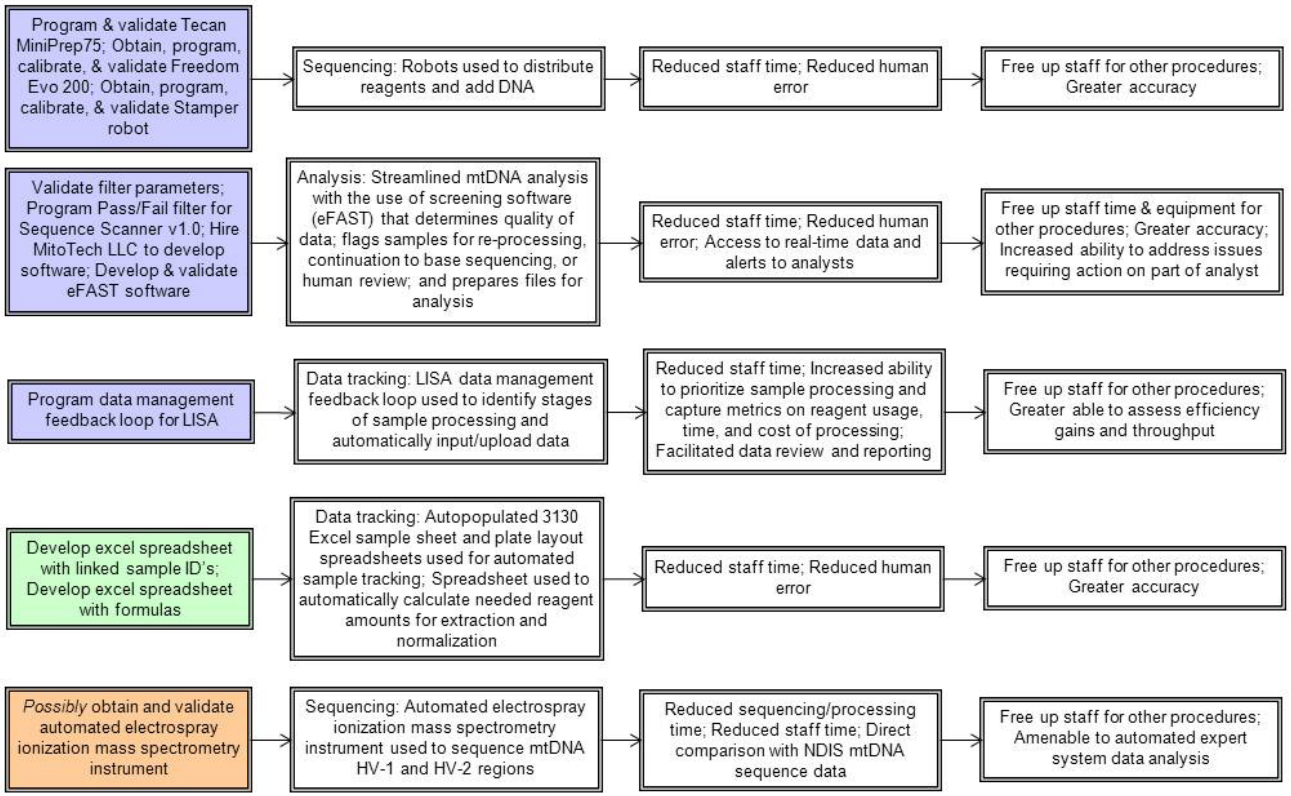
UNT proposed to make a number of changes to the mitochondrial DNA amplification process. Typically, analysts amplify two regions of mitochondrial DNA to be sequenced, the HV-1 and HV-2 sections of the D-loop region. UNT planned to use alternate primers to develop and validate a single amplification procedure that amplified both the HV-1 and HV-2 regions simultaneously rather than amplifying the two regions on separate plates.[62] UNT also proposed optimizing the reagent volumes and number of cycles for the amplification process. The decrease of reaction volumes for single amplifications would result in a reduced need for *ExoSAP-IT* reagents required for template purification. A final step in the proposed changes to the amplification stage was to introduce robotics for automated distribution of reagents and addition of DNA template to the wellplate. These new amplification techniques were expected to lead to a reduction in staff time spent on amplification, plastics and reagents, use of template DNA, human error, and noise in sequence data.

UNT also proposed changes to the quantification and normalization steps of mtDNA processing. UNT proposed to use a *TaqMan* assay to measure the amount of mitochondrial DNA rather than use a nuclear assay, which only provides a rough estimate of the amount of mitochondrial DNA. After doing some research on various assays, UNT decided to partner with the FBI to obtain the necessary chemical reagents and learn about the *TaqMan* test procedure. UNT then planned to validate this procedure within its own lab. The *TaqMan* assay was expected to lead to increased quantification accuracy since it is specifically designed for measuring mitochondrial DNA, which would subsequently lead to more accurate normalization reactions. This could also conserve the amount of template DNA used, reduce the amount of reagents needed, and provide higher quality sequencing data. UNT would then use the *Tecan Freedom EVO 200* robot to automate the normalization process of standardizing the DNA concentration. This use of robotics would lead to reduced staff time spent on liquid handling and a decrease in human error.

[62] And by default, the region in between HV1 and HV2.

Figure 71. Logic Model for UNT Health Sciences Center





The sequencing stages were also expected to be modified through the grant. UNT proposed validating the use of a new sequencing kit (*ABI PRISM BigDye Terminator v1.1 Cycle Sequencing Kit*) and buffer (*The Gel Company's BetterBuffer*) for the cycle sequencing reaction. In addition, during validation, UNT also decided to work on developing an alternate primer that binds in a different location from the D2 primer in order to avoid binding problems that frequently occurred due to a common mutation in the mitochondrial DNA sequence. These combined chemistry changes were expected to lead to a reduction in reagent usage and higher quality sequencing data. UNT also planned to introduce a new bead purification process, the *BigDye Xterminator Purification System*, to remove unincorporated materials (e.g., buffer, primers, dNTPs) in place of the *Performa* columns and plates purification method used previously. UNT expected the modified purification procedures to result in a reduced number of centrifugation steps and transfers of sequencing products from plate to plate, leading to decreased human error from fewer transfers. Further, the new purification system was more amenable to automation.

UNT proposed the use of robotics in multiple capacities during the sequencing stage as well. They proposed to obtain and validate a "stamper" robot that could aspirate the purification beads and rearray the samples from 96 well plates to 384 well plates. UNT also proposed to use robotics to distribute reagents and add DNA template to the wellplates in preparation for capillary electrophoresis. The previously described stamper robot would then add DNA template to the plates. The proposed robotics were expected to reduce the amount of staff time spent in clean-up procedures post-purification and to decrease human error. In addition, more samples could be processed simultaneously with the larger well plates, increasing throughput. Finally, UNT reported in their proposal that they might be able to secure an automated electrospray ionization-mass spectrometry instrument (ESI-MS) from the FBI in partnership with *Ibis Biosciences Inc*. for sequencing the HV-1 and HV-2 regions of mitochondrial DNA. ESI-MS is an instrumental technique that creates electrically charged droplets that contain fragments of the mtDNA molecule, which are drawn into a mass spectrometer for detection. The fragment mass is displayed graphically and interpreted. Mass spectrometry can unambiguously identify DNA bases (A,T,C,G) based on differences in their mass.[63] While UNT did not promise this would be included in their project, if acquired, they expected this new equipment to lead to reduced sequencing time, reduced staff time, and the ability to directly compare sequence data with NDIS mitochondrial DNA sequence data.

For the analysis and interpretation stage, UNT proposed to develop an expert system for filtering data with *rule firings* and to create "middleware" that can link various equipment and expert systems for more streamlined data analysis. In order to do this, UNT needed to optimize and validate rule firings based on filter parameters using DNA data from a Chilean

[63] More specifically, ESI-MS data differentiates based on the mass-to-charge ratio.

databasing project, program *Sequence Scanner v1.0* to categorize data based on these rule firings, and then hire a vendor to develop "middleware" to automatically move data between the capillary electrophoresis instrument, *Sequence Scanner* software programmed as an expert filtering system, and, eventually, a separate expert analysis system (a goal of future development but not within the scope of this grant). After assessing the needs and challenges of creating the proposed "middleware," UNT decided to instead create new software to perform the data filtering and additional functions. The newly developed software, *eFAST* (*Expert Filter Assessment of Sequence Traces*), would be designed to screen raw sequencing data produced through capillary electrophoresis (CE). The system would automatically import data from the CE instrument and determine the quality of the data based on contiguous read length and trace scores. The software would categorize the data from each sample as a "pass" (it can continue to interpretation), "fail" (it needs to be rerun), or "questionable" (the data need to be reviewed by an analyst to determine whether it is usable or not). The *eFAST* software would then automatically organize the data files (sending "pass" samples to one location and "fail" samples to an archived folder) into external file folders in preparation for an analyst to interpret the data (or for importation into an expert analysis system in the future). There are also email alerts that inform the analyst if a sample or plate has failed after the first of six runs in the CE instrument and once sequencing is complete. These alerts allow analysts to respond immediately to sample plates with failed controls rather than waiting to learn about problems after the sequencing is completed (sequencing in their CE instrument typically takes UNT about 5.5 hours). Other advantages to the *eFAST* system over the *Sequence Scanner* software are the abilities to store customized parameters for multiple primers, track and log analyst actions, designate analyst and administrative-user levels, save projects, import and export data files automatically, and interpret control samples according to different rules than DNA samples. Further, UNT planned to eventually make *eFAST* available to the public through open-source code. UNT expected *eFAST* to be beneficial to the lab by reducing staff time spent on reruns, review of raw data, and manual importing of files. The automated file rulings were also anticipated to decrease human error.

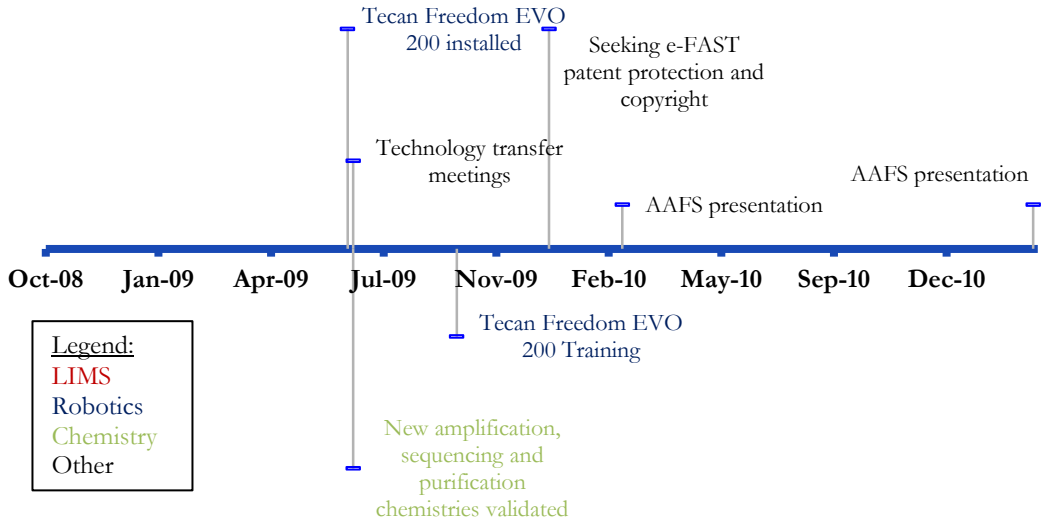
UNT included some modifications to data tracking in their grant goals. A small portion of the grant was set aside for the development of a data management feedback loops in the laboratory's LIMS, *LISA* (*Laboratory Information Systems Application*); these were necessary for the integration of data from the quantification stage, amplification setup, instrumentation and a data flow from expert systems. A more integrated LIMS would reduce time spent on entering data and would allow for more documentation of reagent usage, time, and cost for processing samples. With these additional measures, the lab could more easily evaluate budgetary needs and changes in throughput and efficiency. Autofill *Excel* tools were also planned to automate sample tracking for plate worksheets, calculations for extraction reactions, and calculations for normalization reactions. These tools were anticipated to lead to a reduction in human error and staff time spent manually transferring sample IDs to and

from wellplate worksheets and individually calculating reagent amounts for chemical reactions. These worksheets could be uploaded directly for use by the capillary electrophoresis instrument to track sample layout.

7.3 Implementation Findings

7.3.1 Implementation Description

Figure 72. UNT Implementation Timeline[64]



UNT concluded their implementation on December 31, 2010, 26 months after notification of their award. They received an extension to support a number of implementation challenges, which are discussed below.

UNT's project was extremely comprehensive and included more goals than any of the other sites.[65] UNT accomplished many of the goals set out at the beginning of project. The lab succeeded in using a single amplification of the mitochondrial genome to amplify the HV1 and HV2 regions, automating the amplification setup process, and optimizing the reagent volumes of PCR cycles. The single amplification resulted in a decrease of amplification reagent volumes from 50 µL for the separate amplification of the two regions to 15 µL according to the project director. For quality samples, UNT was able to reduce the

[64] The site was not able to provide additional project milestone dates during the study.

[65] While Louisiana ended up having a similarly large number of goals, these were not set out as such from the start.

number of PCR cycles from 32 cycles to 28 cycles, thus reducing the amount of time the samples undergo amplification. For lower-yield samples, UNT was not able to reduce the number of cycles; however, the reagent volumes were still optimized.

TaqMan, a real-time PCR assay, was validated for the quantification of mtDNA. The validation illustrated that this assay was precise and reproducible even in the presence of inhibitors. UNT also successfully validated the new sequencing chemistries and bead purification system. In addition, UNT designed, ordered, and validated a new primer, which binds in a different location from the usual D2 primer in order to avoid binding problems that frequently occurred due to a common mutation in the mitochondrial DNA sequence. While this was not in the proposal, the development of this new primer occurred as a result of the research into improving sequencing chemistries. Robotics were validated for the cycle sequencing setup; however UNT decided they did not need to rearray the samples since the lab does not receive enough sample requests to make this a useful feature for the lab. UNT was unable to find a stamper robot within the grant's budget that would fulfill the two criteria for aspirating the bead solution: aspirating large beads and small amounts of liquid.

UNT made some additional alterations to their original plan.[66] UNT adjusted its plans in regards to robotics equipment. The proposal included the purchase of a robot for extraction, but UNT instead planned to use these funds for the stamper robot and to assist with the purchase of the *Freedom EVO* instrument. The grant proposal initially requested a purchase of the *Freedom EVO 150*; however, UNT chose to purchase a *Freedom EVO 200* instead since it provided a larger deck platform with more space for samples. The *Freedom EVO 200* was used for an overall validation study of the new research division method. UNT validated the *Freedom EVO 200* for use with normalization, amplification setup for the single amplification region, post-amplification purification, and cycle-sequencing set up. This validation of the robotic platform was conducted using three batches of family reference samples.[67] Before this point, UNT validated the use of robotics more generally, along with the majority of other chemistry changes, with over 1,000 samples on a different robot.

UNT also modified its plans for the software development. They originally proposed to create middleware code to link the *Sequence Scanner* screening software to *Sequencer* analysis software. However, UNT learned from the selected vendor, *MitoTech LLC*, that it would be easier to create new, customized software, which would perform similar screening functions, create additional features, and make data easily available for analysis within a future expert system. UNT created and trademarked a new software product, *eFAST* as described above in section 7.2. By the end of the grant period, UNT had designed and rigorously evaluated *eFAST* v1.1 for accuracy and time savings. As a result, they identified

[66] With approval from NIJ.

[67] This is similar to the timed experiment methodology but was used by UNT for a different purpose.

several ways that the software could be expanded and improved. During the grant period they began work on *eFAST* v2.0, which incorporates additional system rules to more accurately review and sort CE data.[68] Optimization of *eFAST* v2.0 is continuing under a new NIJ award.[69]

UNT acquired and received training on the electrospray ionization mass spectrometer but were unable to begin exploring its capabilities by the end of the grant period. However, this instrumentation was being housed within a different research group at UNT. Finally, UNT accomplished some of the stated goals for the data-management component of the grant. UNT created the Excel worksheets for automated sample tracking and reagent volume calculations. Barcoding software was developed to track samples and reagents. Tracking samples through barcodes, after an initial manual entry of information, could be used to reduce manual data entry during DNA processing and the possibility of transcription errors. Barcoding reagents enables the system to upload reagent information (e.g., lot number, expiration date) to worksheets and alert analysts if the reagent selected has expired before use.

At the completion of this study, the automated sample tracking worksheets were the only outputs of the grant that had been implemented into routine practice in the lab's casework division. Beyond the worksheets, UNT had also proposed to program data management feedback loops in the lab's LIMS, *LISA*, to aid LIMS integration with instrumentation and expert systems. By the end of the grant period, UNT had validated several algorithms within the statistical analysis module of this system. Along with other functions, this module generates random match probabilities, frequencies and likelihood ratios. The status of the proposed feedback loop to aid *LISA* integration was unclear from conversations with the site at the end of the study period.

7.3.2 Implementation Challenges

While UNT moved quickly to initiate grant activities within their research division, they experienced difficulties transferring the validated techniques and equipment into the family reference sample team in the casework division. There did not appear to be a collaborative process between the research division and the casework division teams to identify the lab's greatest needs, resulting in some novel approaches not being accepted and implemented by the casework division. For instance, the *TaqMan* assay, Excel worksheet tool for automated normalization calculations, and use of robotics to automate the quantification and normalization stages were not viewed as especially helpful by the casework division because

[68] The additional sorting rules include high baseline, high signal, low signal, partial read, mixture, homopolymeric stretch, and length heteroplasmy.

[69] Funded by NIJ Award 2009DN-BX-K171.

they did not perform quantification and normalization steps for family reference samples. These samples tended to be of high quality and have similar amounts of DNA.

Other implementation challenges for the grant included some delays in purchasing and validations. UNT did not obtain the main robotics system, the *Freedom EVO 200*, until June 2009. Before that point, the research division staff programmed, calibrated, and validated automation scripts on a different liquid handling robot, the *Tecan MiniPrep 75*. UNT performed in the initial validation of its new process with this earlier robot. Liquid calibration problems with a second *Tecan Miniprep 75* remained unresolved for eight months and prohibited the validation of automated sequencing setup reactions with actual samples. UNT instead established proof of concept through the use of colored dyes after the second *Tecan Miniprep 75* was working again. Once the *Freedom EVO 200* arrived, the research team worked on creating automation scripts, calibrating, and validating the new robot. The robotics vendor spent a month addressing some issues with internal tubing in the *Freedom EVO 200*, causing some delays in this validation work.

UNT also made some changes to the use of permanent and disposable pipette tips when moving from the *Tecan Miniprep 75* to the *Freedom EVO 200*. During the initial validation with the *Tecan Miniprep 75*, UNT used permanent pipette tips with a bleach wash that purified the tips between aspirating liquids. In contrast, for the *Freedom EVO 200*, UNT used a bleach wash to purify the pipettes holding controls and reagents, but used disposable tips when DNA was aspirated. This helped reduce the robotics time further since less time was needed for the robot to purify its pipette tips in bleach.

7.3.3 Final Perceptions

The research team spoke with the key contacts for the grant project in the research division and the casework division at the end of the study period. Both interviewees conceded that the full effects of the grant's innovations had not been felt yet. The manager of the grant project viewed the grant as having important outcomes for the lab that could not be fully realized until the submissions input of family reference samples increased. Because the lab did not have a backlog or receive a continuous and high-quantity stream of samples, it would be difficult to implement all of the grant interventions (which are designed for a high-throughput workflow). However, the new process had been developed, validated, and was ready for implementation once the need was present. The interviewee thought more educational work was needed to inform and train DNA collection agencies about the role of DNA in missing persons' identification (which should subsequently increase their family reference sample submissions). While the casework division leader also felt that little had changed in terms of actual casework due to the grant, both individuals reported a positive outlook toward additional implementation in the future. At the time of the interview, the

casework division was validating the single amplification within its current workflow, and there was strong interest in the robotics and *eFAST* system for the future.

Both individuals also reported important strides being made during the project in terms of improving communication and facilitating an exchange of ideas between the research and casework arms of the laboratory. This project brought awareness to the pitfalls of isolating the two sections from each other, and efforts were being made to facilitate better collaboration in the future. For instance, the two divisions of the lab now meet regularly to discuss goals and updates of their respective work.

Both interviewees reported that the biggest challenges of the project were implementing the new procedures into the casework division, primarily due to a lack of communication and mutual engagement during project planning. Other difficulties mentioned were securing the 25 percent match and completing tasks within the grant's original 18-month period of performance. The casework division interviewee also said that their division would be more comfortable making incremental modifications rather than wholesale changes of the entire process.

Overall, both representatives identified some important lessons learned from the current project. While the majority of the grant's components have not been implemented into actual casework at UNT, both divisions were actively communicating and working together toward future implementation of those pieces viewed as most beneficial by both sides. Finally, the head of the grant project expressed disappointment that the NIJ efficiency initiative had ended, believing that efficiency gains were an important goal and should continue to be pursued.

7.4 Outcome Findings

The following section describes the data used to assess the outcomes of the NIJ Forensic DNA Unit Efficiency Improvement Program on UNT and changes in productivity and efficiency at UNT.

7.4.1 Descriptive Statistics

During the period between January 1, 2010, and February 28, 2011, the UNT laboratory analyzed 3,428 family reference sample "routings,"[70] 79 percent of which involved mitochondrial DNA analysis. These routings were related to 3,136 laboratory cases. Cases had up to four routings per case, although the majority of cases (89.8 percent) had a single routing. Routings had a range of 1–8 family reference samples each, with an average of 1.47

[70] A routing was defined by the site as any time a sample or set of samples underwent the DNA process. This construct is somewhat comparable to the idea of a submission or request.

samples per routing. A small proportion of the routings were considered pending[71] (0.5 percent), on hold (0.1 percent), or were canceled by the lab (1.2 percent).

Offense information is not known, because family reference samples are used for missing person cases. Until a match is made to either a living person or human remains, nothing is known about the type of crime or whether one was committed at all. The majority (69.7 percent) of routings had the medium priority level, over one-quarter (27.2 percent) had the highest priority, and only 3.1 percent had a lower priority level. Twelve analysts were listed as responsible for the selected routings.

The UNT laboratory completed about 56 (median) routings per month across the four years (5.6 routings per month per analyst) (see table 11). The median turnaround time was 64 days from the start of testing until completion of the technical review. Three stages were available for analysis in the study: (1) date started until extraction completion, (2) extraction completion until interpretation completion, and (3) interpretation completion until technical review completion. The first stage was fairly brief with a median of 2 days between start and extraction, while the two latter stages had medians of 22 days and 29 days, respectively.

[71] Cases may have a "testing pending" status if they are waiting on additional samples for an incomplete submission or are in analysis queue waiting to be assigned.

Table 11. University of North Texas Throughput and Turnaround Time Outcomes

Family Reference Samples (*N* = 3,428)		Productivity/Labor	Cleaned Productivity	Raw Productivity
Overall Outcomes				
Routing Turnaround Time	Mean	5.84	60.04	59.78
Start–Technical Review	Median	6.30	64.00	64.00
	Std. Dev.	2.60	26.81	68.81
	Range	(.10, 16.20)	(1, 162)	(-3178, 1337)
Routing Throughput	Mean	6.41	N/A	65.75
	Median	5.60	N/A	56.00
	Std. Dev.	3.73	N/A	38.36
	Range	(0.18, 15.10)	N/A	(2, 155)
Stage-Level Turnaround Time				
Date Started–Extraction	Mean	0.35	3.60	3.57
	Median	0.20	2.00	2.00
	Std. Dev.	0.39	3.93	9.08
	Range	(0, 2.80)	(0, 28)	(-365, 298)
Extraction–Interpretation	Mean	2.40	24.81	28.51
	Median	2.20	22.00	22.00
	Std. Dev.	1.79	18.39	135.21
	Range	(.09, 42.70)	(1, 427)	(1, 7351)
Interpretation–Tech Review	Mean	3.10	31.85	27.92
	Median	2.80	29.00	29.00
	Std. Dev.	2.12	21.68	144.13
	Range	(0, 12.20)	(0, 122)	(-7273, 471)

Notes: Labor is defined as the number of staff reported for that year. Budget is defined as the annual DNA unit budget in $100,000 units. Turnaround time is reported in number of days. The research team analyzed the stage of interpretation both for all data interpretation and for mitochondrial DNA interpretation on its own. Results were nearly identical between the two for descriptive statistics as well as regression findings. Therefore, for the sake of simplicity and because the implemented intervention is expected to affect **all** family reference samples, the report only presents findings from the overall data interpretation rather than the mitochondrial DNA-specific data interpretation.

Across the four years of data, there is substantial variability in all of the outcome measures (see figures 73–83). The black, vertical line delineates when the autofill worksheets began being used by the FRS staff. The absence of a stable baseline does not allow for easy interpretation. However, in general, there does not appear to be any strong pattern of findings suggestive of an impact of the grant program on either the number of routings started[72] (figure 73), throughput (figure 74), or turnaround time (figure 76). The only distinguishable trend is that there was a multi-month reduction in turnaround time in late 2007–mid-2008. The lab did not report any events to the research team, which might explain this change. Efficiency measures of throughput (figure 75) and overall turnaround time (figure 77) also do not reveal any strong patterns; however, 2009 appears to be the most efficient year when labor resources are taken into account. This is likely not due to the grant itself, given the minimal implementation and absence of effects seen in productivity measures.

No clear patterns were detected in any of the three stage-level turnaround time measures. The time between starting a case and completing extraction was typically short, although there were three large spikes in turnaround time during the study period (figure 78). The turnaround time between extraction and interpretation may have experienced an increase around the midpoint of the study, although is difficult to determine with the many data spikes present (figure 80). The time between interpretation and technical review showed a similar pattern to that of the overall turnaround time measure, with a dip in turnaround time beginning in fall 2007 (figure 82). Efficiency estimates varied by stage on which year was most efficient in terms of labor resources (see figures 79, 81, and 83). Because no consistent pattern was detected among the performance measures and the lab reported no other significant changes occurring at the laboratory during the study period, it is likely that the changes seen are due to random variability in casework.

[72] Completed cases are not necessarily the same cases as those started each month. Started cases are matched to the month in which a case was "started," while completed cases are assigned to the month in which the case was completed.

Figure 73. Monthly Number of Routings Started

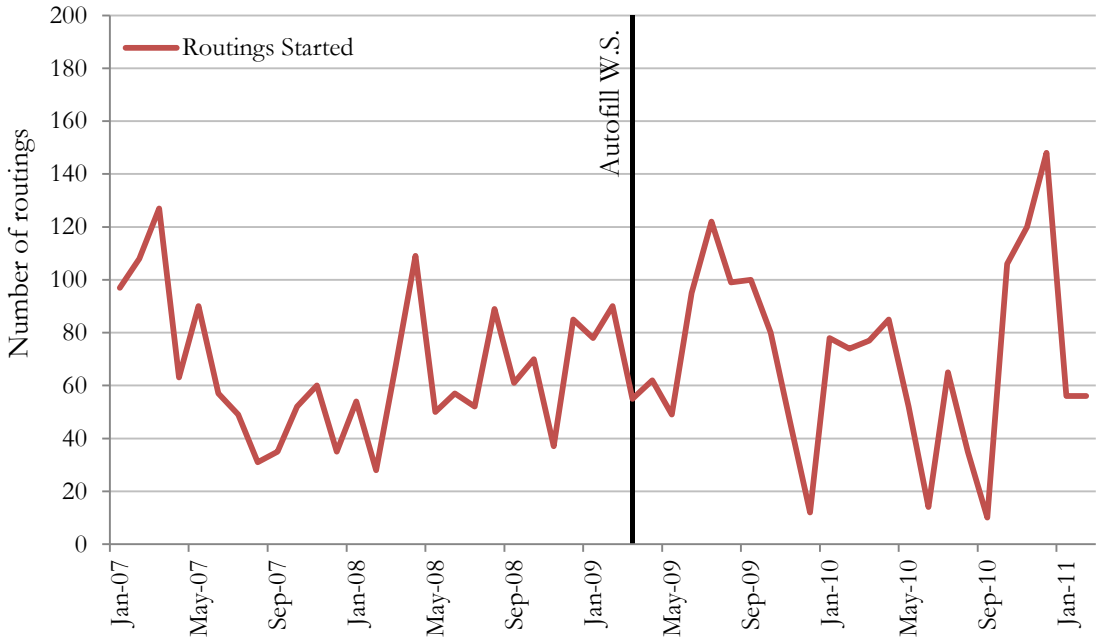


Figure 74. Monthly Throughput of Routings

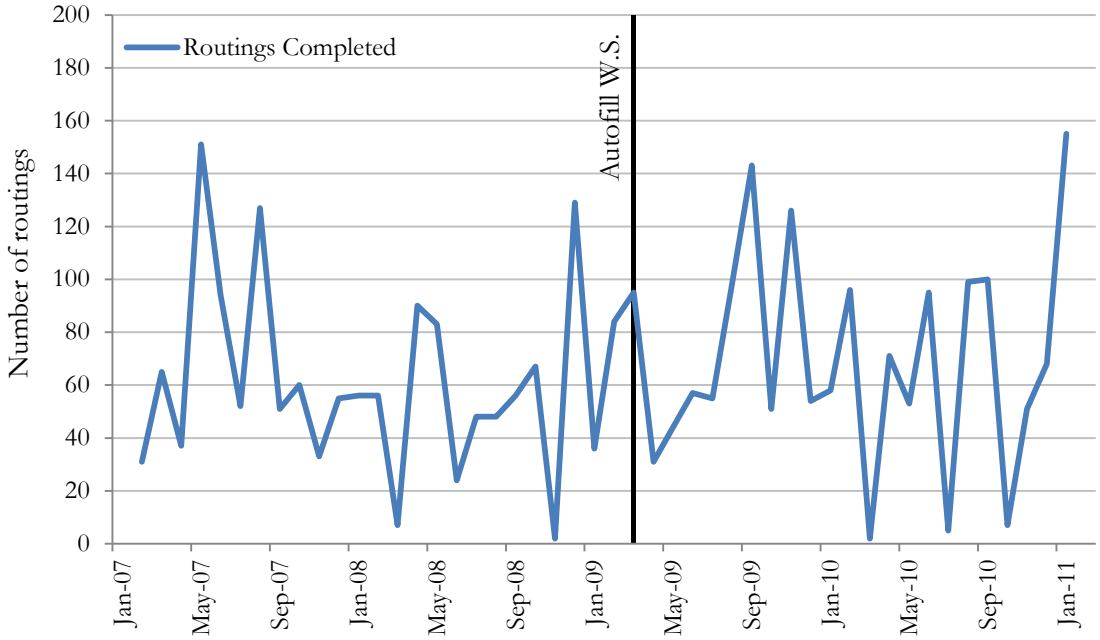


Figure 75. Efficiency Measure of Monthly Throughput by Labor

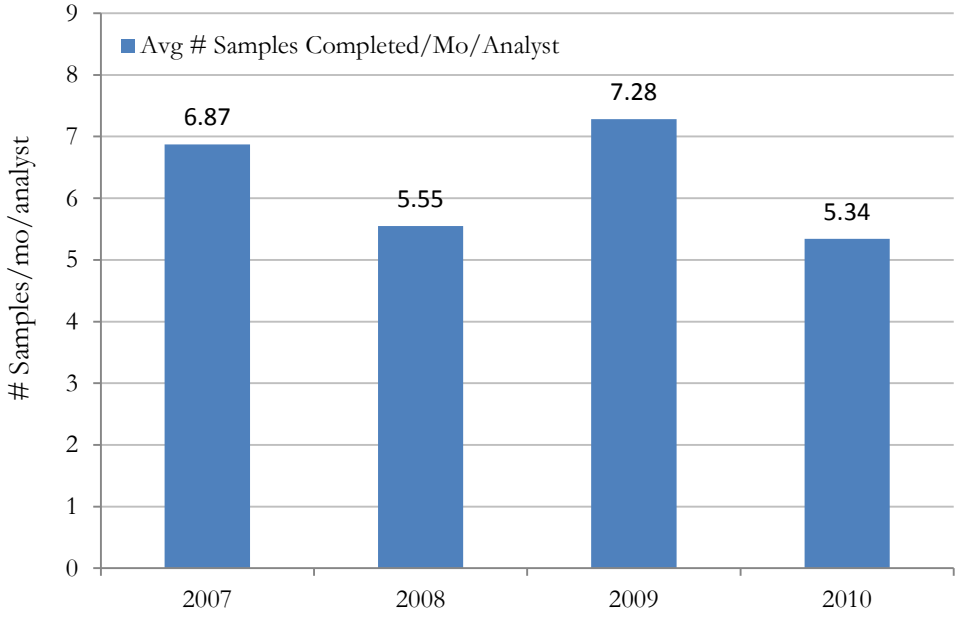


Figure 76. Routing Turnaround Time

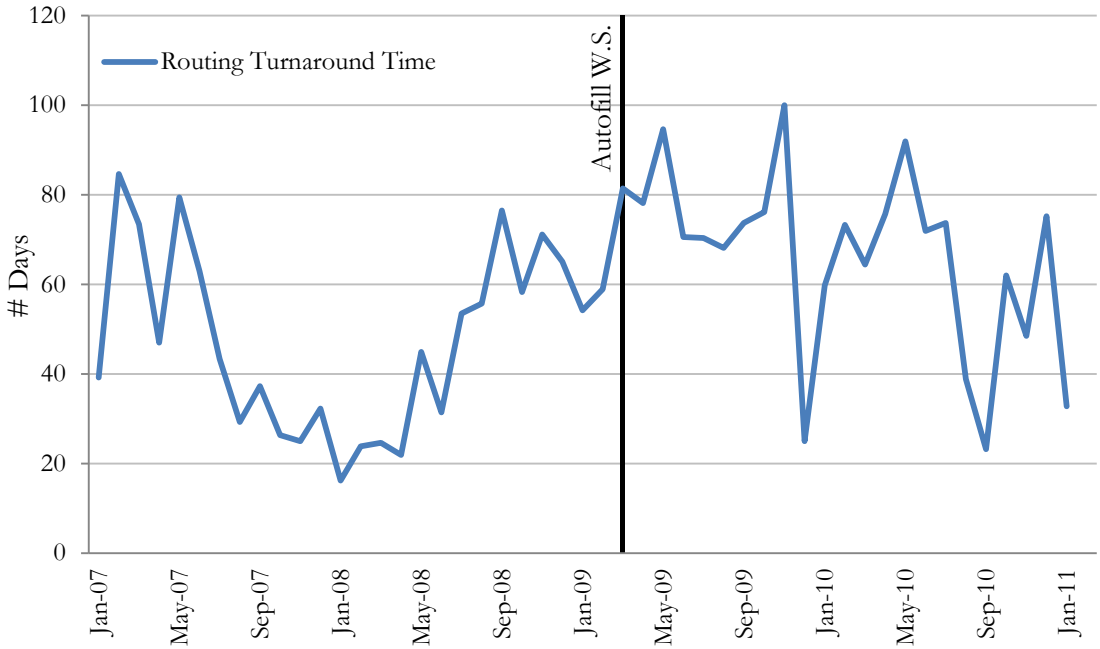


Figure 77. Efficiency Measure of Turnaround Time by Labor

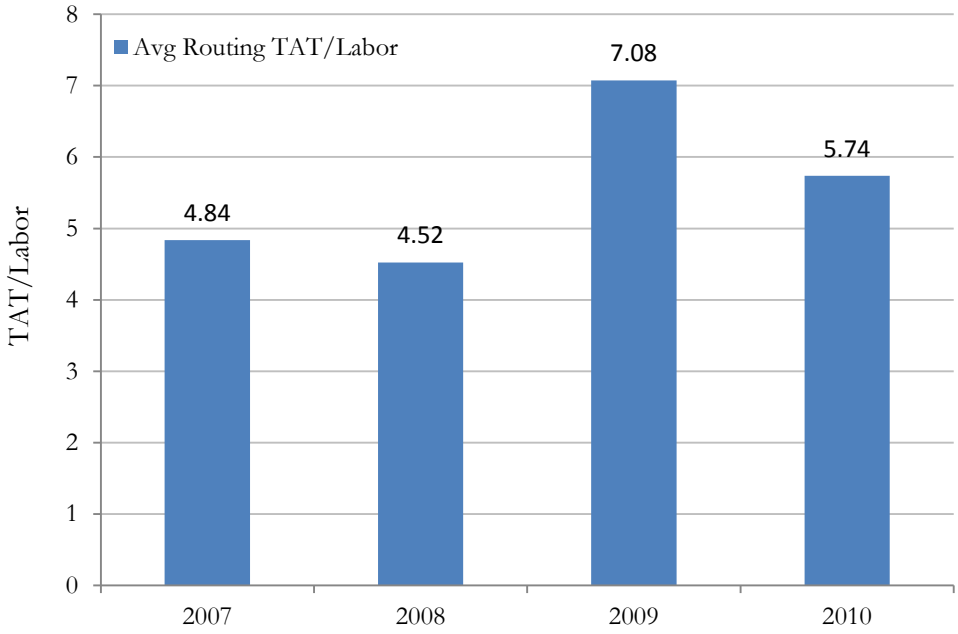


Figure 78. Stage-Level Turnaround Time: Start Date to Extraction

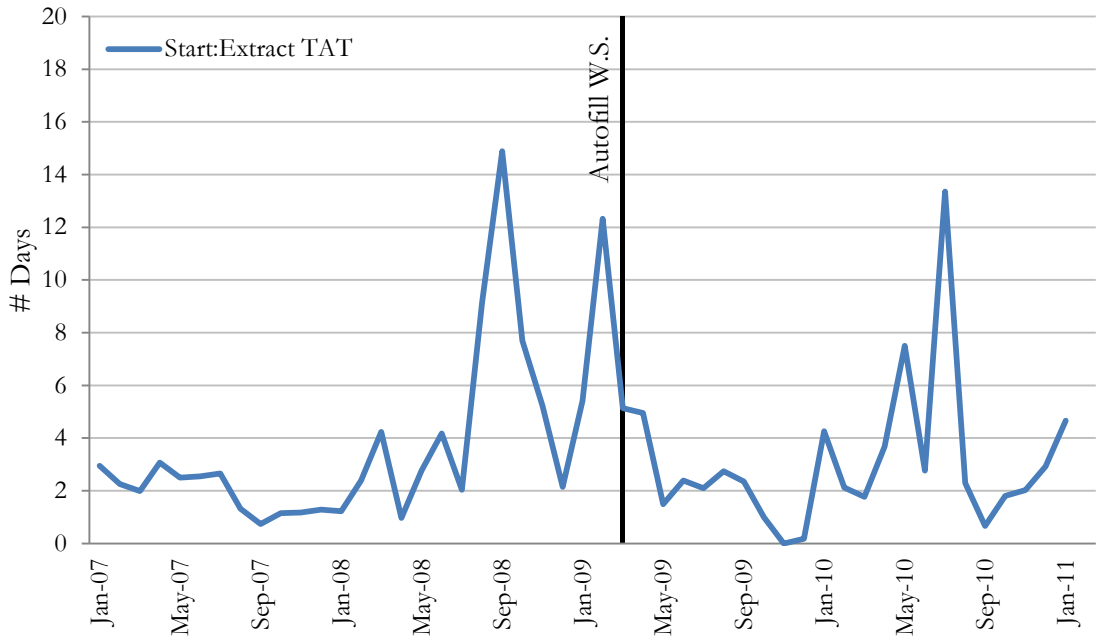


Figure 79. Stage-Level Efficiency Measure by Labor: Start Date to Extraction

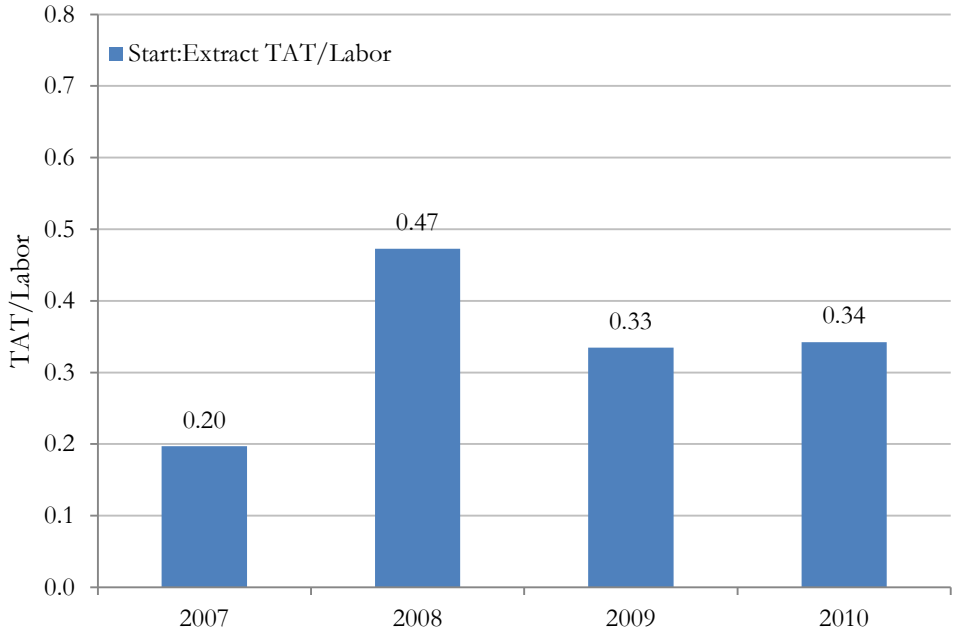


Figure 80. Stage-Level Turnaround Time: Extraction to Interpretation

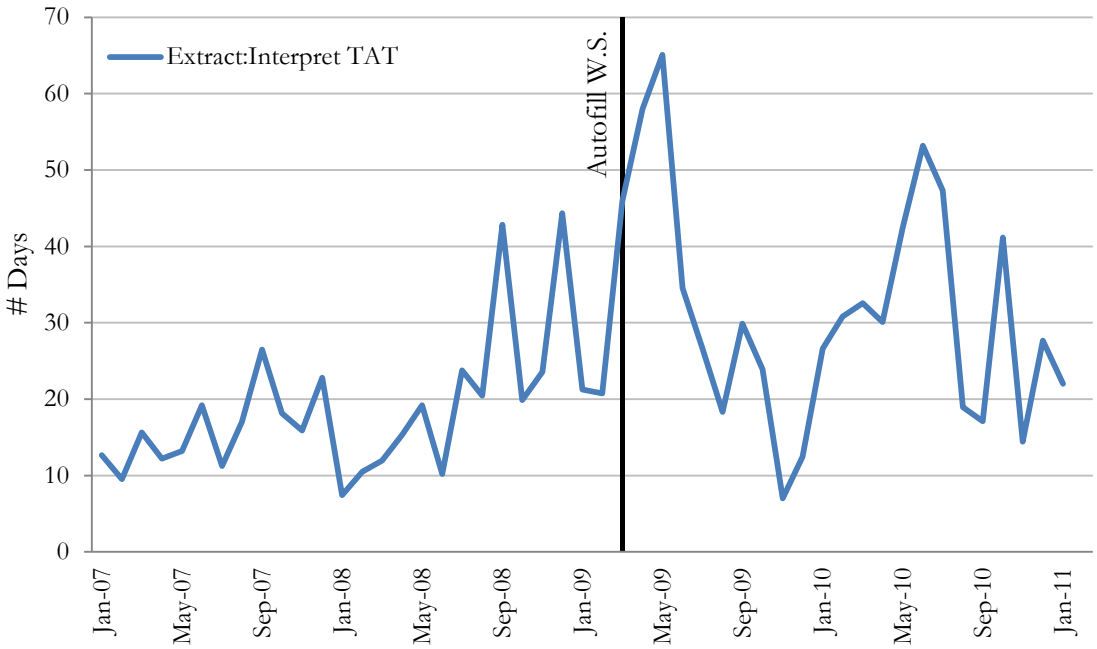


Figure 81. Stage-Level Efficiency Measure by Labor: Extraction to Interpretation

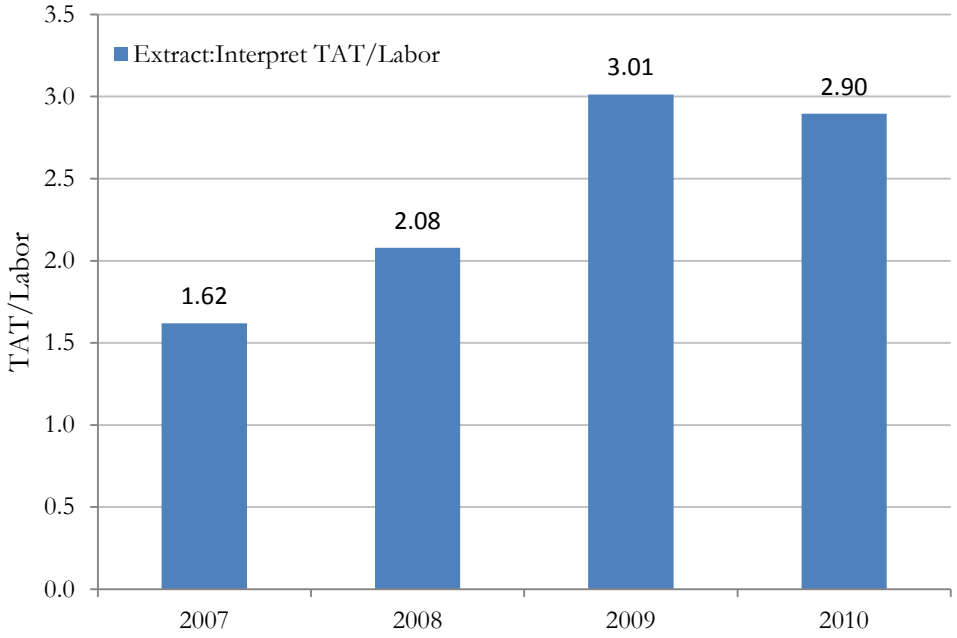


Figure 82. Stage-Level Turnaround Time: Interpretation to Technical Review

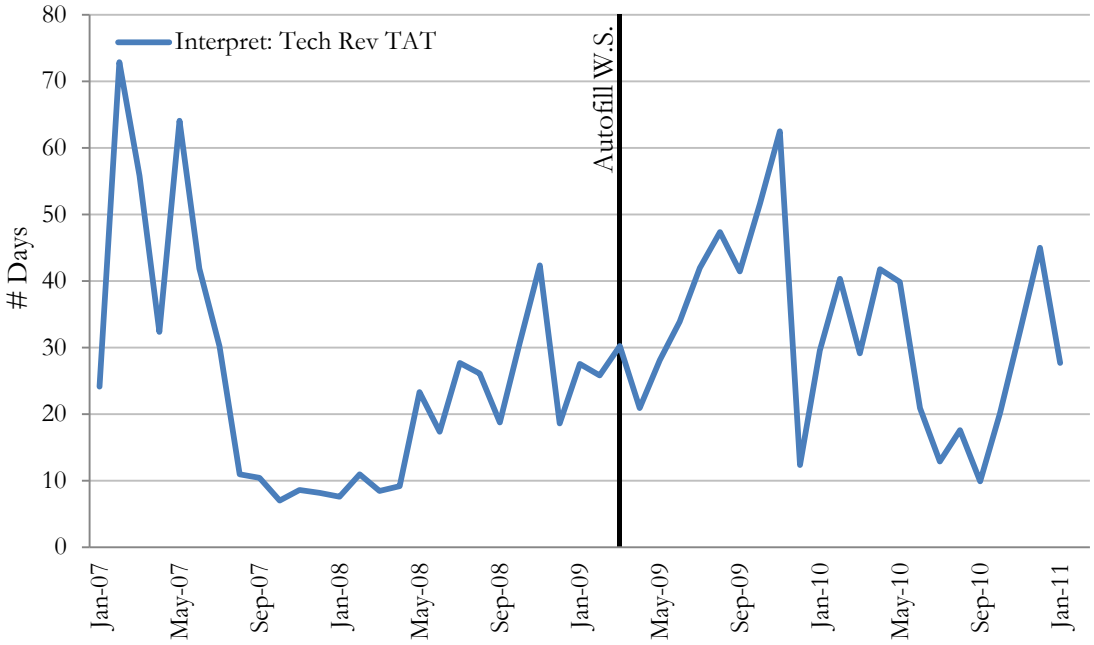
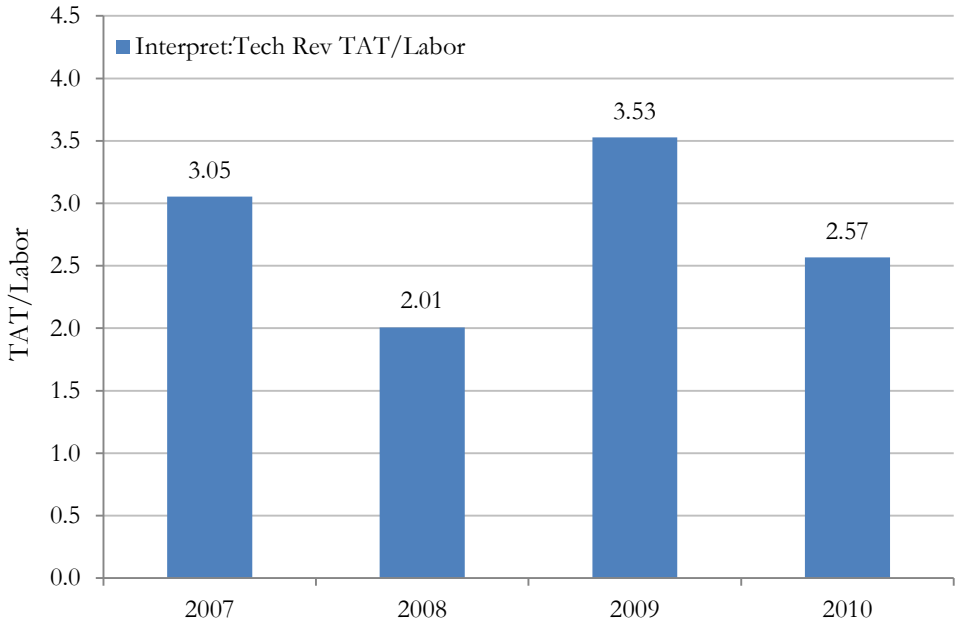


Figure 83. Stage-Level Efficiency Measure by Labor: Interpretation to Technical Review



7.4.2 Pre/Post Comparisons

The research team conducted two types of tests (independent samples t-test and Mann-Whitney U test) to compare the throughput both before and after the single implementation milestone: the use of automated sample tracking worksheets. Both the independent samples t-test and Mann-Whitney U test are used to compare differences between two groups. However, the Mann-Whitney U uses the median as the measure of central tendency and, therefore, does not require normality as the independent samples t-test does. Although t-tests are robust to violations of the normality assumption, the Mann-Whitney U test was also conducted since the data were skewed.

No significant change in throughput was observed when comparing the number of routings completed per month before and after implementation of the automated worksheets (see table 12). This was true regardless of test used and for both the pure productivity measure and the efficiency measure of monthly throughput divided by the number of staff. Therefore, it does not appear that the grant had an impact on overall throughput for the UNT laboratory.

Table 12. Comparison Test Results for University of North Texas

Throughput	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of automated worksheets (Mar. 2009)	-0.76	0.45	240.00	0.33
Throughput/Labor	**t-test**		**Mann-Whitney U test**	
	t statistic	*p-value*	*U statistic*	*p-value*
Implementation of automated worksheets (Mar. 2009)	-0.42	0.68	264.50	0.63

7.4.3 Regression Analyses

In order to detect whether the program had an effect on turnaround time and its related efficiency measures after controlling for other case characteristics, the research team performed a series of negative binomial regression analyses (see table 13). Regressions were used to model overall and stage-level turnaround times, both as pure productivity measures (case turnaround time) and as efficiency measures (case turnaround time divided by the number of staff during the year the case began). The research team included a dummy variable that indicated whether the case occurred after the implementation of the automated worksheets in March 2009. The site did not report any other changes or events in the lab expected to influence turnaround time (other than staff changes which should be reflected in the efficiency estimate).

In contrast to the t-test and Mann-Whitney U findings, the analyses revealed a significant effect of the March 2009 implementation milestone. However, the direction of the coefficient was positive, indicating that cases processed after March 2009 had longer turnaround times than those processed before that point. There is no theoretical reason the automated worksheets should increase the overall and two of the stage-level turnaround time measures; therefore, it is likely that this trend is coincidental and not related to the grant program. However, it is further evidence that the automated worksheets did not improve turnaround time at the laboratory. On the other hand, the time between when DNA work was started until extraction was completed appeared to decrease after March 2009.[73]

However, the lab reported that the automated worksheets should affect many stages throughout the process and not just this first stage. Therefore, it is unclear whether this decrease in turnaround time can be attributed to the grant.

[73] The lab defines this "start" date loosely, but it seems to most closely align to when the analyst first comes into contact with the evidence materials.

The regressions revealed that turnaround time varied by other influential case characteristics. The priority level of a case was significantly related to the overall turnaround time, as well as the stage-level turnaround times between the start date and extraction, and between interpretation and technical review. Routings with lower priority (rated 3 on a 1–3 scale) tended to take longer during these stages, while priority had no influence on the time between extraction and interpretation. The opposite pattern occurred with the variable indicating whether the routing required mitochondrial DNA analysis. Routings with mitochondrial DNA had longer turnaround times between extraction and interpretation, although there was no effect of mitochondrial DNA at the beginning or end stages. This aligns with what the researchers heard from the site that mitochondrial DNA takes significantly longer time to analyze due to its method of amplification and data interpretation. This increase could also be seen at the level of overall turnaround time. Analyst experience did not affect the overall turnaround time of a routing; however, it was related to the three stage-level turnaround time measures. More experienced analysts tended to take more time processing family reference samples from start to interpretation (possibly due to competing management or supervision obligations). However, they spent less time during the report and review stages. Another variable, the number of samples, could not be included in the regression, because it reduced the sample size too much due to a high number of missing values. However, in separate analyses which included this measure, the number of samples was generally unrelated to turnaround time for routings (which makes sense since UNT uses batch processing).

Efficiency measures of routing turnaround time divided by annual labor numbers mirrored the turnaround time outcome findings, with the exception that analyst experience became significant for overall routing turnaround time divided by labor.

Table 13. Regression Results for University of North Texas

Overall TAT Regression	**Overall Routing TAT**		**Overall Routing TAT/Labor**			
	b coeff.	*p-value*	*b coeff.*	*p-value*		
Intervention: Autofill worksheets	0.36	<.01	0.32	<.01		
Priority	0.15	<.01	0.13	<.01		
Mitochondrial analysis	0.12	<.01	0.12	<.01		
Analyst Experience	0.01	0.27	-0.01	0.01		
Stage-Level Regression	**Start–Extract TAT**		**Extract–Interpr TAT**		**Interpr–Tech Rev TAT**	
	b coeff.	*p-value*	*b coeff.*	*p-value*	*b coeff.*	*p-value*
Intervention: Autofill worksheets	-0.28	<.01	0.52	<.01	0.30	<.01
Priority	0.24	<.01	-0.003	0.86	0.30	<.01
Mitochondrial analysis	0.08	0.09	0.25	<.01	0.02	0.58
Analyst Experience	0.11	<.01	0.04	<.01	-0.03	<.01
Stage-Level Regression	**Start–Extract TAT /Labor**		**Extract–Interpr TAT /Labor**		**Interpr–Tech Rev TAT /Labor**	
	b coeff.	*p-value*	*b coeff.*	*p-value*	*b coeff.*	*p-value*
Intervention: Autofill worksheets	-0.35	<.01	0.48	<.01	0.26	<.01
Priority	0.19	<.01	-0.02	0.43	0.28	<.01
Mitochondrial analysis	0.07	0.42	0.23	<.01	0.02	0.63
Analyst Experience	0.09	<.01	0.02	<.01	-0.04	<.01

7.4.4 UNT's Timed Experiment

UNT was unique in this study in that by the end of the grant period, a new DNA processing method was operational in the research division that had not been implemented into the casework division.[74] As a result, the case processing data from the casework division could not be used to evaluate the operational interventions that were not implemented. However, as described in section 2.4, UNT Timed Experiment Supplemental Analysis, this site performed a timed experiment with three batches of family reference samples. Results for the UNT timed experiment are shown in the table 14 below.

Table 14. Results of UNT Timed Experiment: Comparison of Casework Family Reference Sample and Research Division Methods

DNA Processing Stage	DNA Processing Substage	Research Division Method Average Time (min)	Family Reference Sample Method Average Time (min)	Stage-level differences (FRS method average – RD Method average) (min)
Quantification	Consumables and Master Mix Prep	15.2	13.7	-1.5
	Plate Prep	15.2	20	4.8
	Quantification - 7500	99.0	90	-9.0
Normalization	Consumables and Reagent Prep	12.0	0	-12.0
	Worksheet Prep	19.8	0	-19.8
	*Plate Prep	12.3	0	-12.3
Amplification	Identifiler Consumables Prep	12.0	15.7	3.7
	Identifiler Master Mix Prep	7.2	5	-2.2
	*Identifiler Plate Prep	7.9	16.4	8.6
	Identifiler Amplification	186.7	188	1.3
	mtDNA Consumables Prep	12.3	30	17.8
	mtDNA Master Mix Prep	11.8	11.5	-0.2
	*mtDNA amp Plate(s) Prep	7.0	15.5	8.5
	Mito Amplification	106.0	104	-2.0
Mito Sequencing	Agilent Reagent Prep	0.0	30	30.0
	Agilent Prep and Run	0.0	45	45.0
	ExoSAP-IT Reagent Prep	2.7	1	-1.7
	ExoSAP-IT Addition	10.5	9.1	-1.4
	ExoSAP-IT	34.7	33	-1.7

[74] The Excel autofill worksheets were implemented by the casework division and used by that division in this timed experiment.

DNA Processing Stage	DNA Processing Substage	Research Division Method Average Time (min)	Family Reference Sample Method Average Time (min)	Stage-level differences (FRS method average – RD Method average) (min)
	Cycle Sequencing Master Mix Prep	5.0	14.0	9.0
	Plate Prep	13.8	27.5	13.8
	Cycle Sequencing	127.8	80.0	-47.8
	Reagent/Column Prep	10.0	30	20.0
	Post CS Clean-Up	30.0	17.5	-12.5
Mito Analysis	Initial QC Review and Base Calls	130.0	705	575.0
IdentiFiler	Reagent Prep	4.4	5	0.6
	*Plate Prep	12.8	25.2	12.4
STR Analysis	Initial QC Review Only	20.2	45	24.8
		Total time difference (min):		**651.1**
		Total time difference (hrs):		**11**

* Substages were performed robotically by the research division and manually by the casework division. There may be additional analyst time savings on these steps due to the lack of supervision needed for robotic technologies.

Average DNA processing substage time was calculated from the individual times self-reported by the site. Differences between these averages illustrated the time savings (shown in green) or time losses (shown in red) associated with the research division DNA processing method.[75] The time for the electrophoresis subtask is not included in this comparison. Since the batches being compared are not compositionally identical (i.e., they contain different numbers of different family reference samples), it cannot be known if any time differences observed in the electrophoresis substage are due to differences between the methods or are attributable to differences in batch composition (i.e., number of samples). Additionally, the average times shown for the STR analysis substage represent the initial QC review and do not include the complete STR data analysis; this was necessary because the research division time data reported was only for this first review.

Overall, and for the stages shown in table 14, the new research division process had an average time savings of 651.1 minutes (or 11 hours) for each batch when compared to the casework division family reference sample method. The vast majority (88%) of this time savings was obtained during the mtDNA data analysis step. This includes the initial QC

[75] The time for the electrophoresis subtask is not included in this comparison because that time is solely dependent on the number of plates, which is a function of the number of samples in each batch and is also affected by division policy. It was reported to the evaluators that as a result of casework policy, only a portion of the samples in each batch undergo mtDNA processing. The R&D section processed all samples in the batch for mtDNA, therefore the number of plates and the electrophoresis times are influenced by policy as well as differences in laboratory procedures.

review and all base calls to determine the mtDNA sequence. Total times for each DNA processing method are not presented because the sum of these averages does not represent the total DNA processing time for each method.[76] The data reported by the sites represent minutes that the analyst spent actively working on substage tasks and time the batches spent being manipulated by instrumentation. Times between the substages were not recorded by the casework section and therefore these values cannot be summed to provide a total method turnaround time. However, the total time differences can be calculated because this is a function of the individual substage times in table 14.

It should be noted that the time savings shown are averages and that each DNA processing step had a range of values from three to eleven. As a result, caution must be taken when interpreting these results. What does seems clear is that, for these three batches,[77] the robotic technologies coupled with the *eFAST* data sorting yielded a time savings over the business-as-usual processes of the FRS section especially in the data analysis step. These timed data collections and calculations could be repeated once these interventions are implemented into active casework for a larger number of batches to produce more generalizable results. Negative values, shown in red, indicate DNA processing stage subtasks where the casework family reference sample method took less time than the research division method. Positive values, shown in green, identify stages where the new research division method yielded a time savings. The time loss during the normalization stage is due to the fact that the casework division does not perform those tasks on family reference samples.

In addition to the time data recorded, and for the batches both tested in each division, the research division supplied quality data for STR analysis including first-pass allele count.[78] This shows how many alleles were detected from the samples in each batch and can serve as a proxy for method success when sample quality and quantity are held constant. If alleles were missed in the first run, DNA processing stages may have been repeated.[79] Table 15 shows the allele call comparison between the identical batches processed by the research division method and the casework family reference sample method and the expected number of alleles expected.[80] In each of the three batches, DNA processing data from the research division method called more alleles and therefore resulted in fewer missing alleles than the casework family reference sample method. Due to the low number of observations (n=3), care must be taken when trying to generalize from these results. However, these results show

[76] Due to the purposely omitted electrophoresis stage and the lack of data on the time in-between each stage.

[77] For some substages, there were additional data points due to time data from dye tests and cancelled batches.

[78] We must note that these are self-reported data, not observations made by the evaluation researchers.

[79] Alleles are missed if they are present below the instruments level of detection (LOD) or if they are detected but below the intensity minimum threshold to be called.

[80] Unlike the time data, these quality data are for the same three batches so a direct comparison in more valid.

improvements for these three batches, and a similar comparison should be made with more observations.

Table 15. Comparison of the Research Division and Casework Family Reference Sample Methods: Missing Alleles from Three Batches

	Batch A			Batch B			Batch C			
	Expected: 2,314			Expected: 2,375			Expected: 2,182			
	Detected	Missing		Detected	Missing		Detected	Missing		Total Missing
RD Method	2,314	0		2,366	9		2,182	0		9
FRS Method	2,269	45		2,347	28		2,057	125		198

7.4.5 Conclusions

The Center for Human Identification within the University of North Texas Health Sciences Center at Fort Worth proposed to implement a series of new approaches to analyzing mitochondrial DNA family reference samples, including changes related to chemistry, robotics, expert filtering software, and data tracking. UNT's Field Testing Division developed and validated these techniques for eventual implementation by the family reference sample team of the forensic casework division. While UNT moved quickly to initiate grant activities within their field testing division, they experienced difficulties transferring the validated techniques and equipment into the family reference sample casework team. Of the many ambitious tasks set out by the grant (nearly all successfully validated), only the autofill Excel worksheets were implemented into routine casework. These implementation challenges were primarily due to the fact that the grant's effects could not be fully realized until the lab had a sustained stream of sample submissions, as well a lack of mutual engagement and collaboration between the casework and research divisions. While the majority of the grant's components had not been implemented into actual casework by the end of the study period, both divisions were actively communicating and working together towards future implementation of those pieces viewed as most beneficial by both sides.

There was substantial variability in productivity outcome measures (i.e., throughput and turnaround time) across both cases and time which created additional challenges in detecting patterns. From visual inspection of the graphs, the data do not suggest a strong impact of the grant program on either the throughput or turnaround time (global or stage-level). Efficiency measures of throughput or overall and stage-level turnaround time also do not reveal any strong or consistent patterns. Similarly, no effect was detected by pre/post comparison tests or regression analyses. However, it is possible that other benefits, such as greater accuracy from automatic entry (as opposed to manual entry which can result in human error), may have resulted from the implementation but cannot be observed from this study's analysis. Additional characteristics that were related to longer turnaround times were having a lower

priority level, needing mitochondrial DNA analysis, and being assigned to a more experienced analyst (although these analysts spent less time during the report and review stages).

The timed experiment illustrates that for some DNA processing stages, the casework family reference sample method takes the same or less time than the research division method. However, the large difference in time needed for analysis of mtDNA sequence data indicates that the eFAST software could generate considerable time savings over the analysis method currently used by casework division. This result should be confirmed in a future data collection after multiple sites have adopted this filtering software.

Although the current evaluation could not detect effects of the grant by the end of the study period, this may be due primarily to a lack of implementation and *not* to the ineffectiveness of the proposed approach. UNT had a very comprehensive and ambitious strategy, and they were able to successfully develop and validate the majority of their goals. Based on the site's own timed experiment, there is strong potential for substantial time savings gains with the newly developed workflow. However, until more of the grant outputs are implemented into casework, no conclusion can be drawn on the actual outcomes of such an approach on family reference sample throughput and turnaround time.

8. CROSS-SITE FINDINGS AND IMPLICATIONS

This concluding section of the 2008 NIJ Forensic DNA Unit Efficiency Improvement Program evaluation report summarizes cross-site implementation findings and lessons learned from efforts to improve crime laboratory DNA processing performance through novel innovations designed to increase efficiency instead of capacity. It also presents select outcome findings that might be associated with the processes, activities, technologies, and other approaches in field settings under grant funding by NIJ.

In general, a wide range of promising and viable approaches to improve DNA evidence processing were proposed across the sites, although two of the six original grant recipients withdrew from the program.[81] The initiatives at each site were the result of considerable local level effort supplemented by guidance and support from NIJ. However, the sites also faced significant implementation challenges. Measurement of program outcomes was also a challenge given the evidence processing data routinely collected and maintained by publicly funded crime labs. Overall, outcomes were mixed both within and across sites. However, the research evidence compiled for this evaluation does provide some support for the hypothesis that crime laboratories can effectively implement successful methods to improve efficiency.

8.1 Implementation Lessons Learned

8.1.1 Implementation Delays

Past research and experience associated with the implementation of new and innovative organizational policies and practices has clearly documented the difficulties that often arise in spite of the best of intentions. Unanticipated implementation delays often occur, and full implementation of program plans can be affected accordingly. There are a variety of explanations for such implementation challenges, but the fundamental reason is that demonstrations of new approaches take place in the real-world settings of organizations, not in controlled scientific settings. As a result, day-to-day workload demands, internal resources (personnel availability and skills levels), and external environmental constraints (legal and policy) can unexpectedly impede implementation efforts.

It is therefore not surprising that there were significant implementation delays across the study sites. First, crime laboratories across the United States, including those that participated in this program, have been faced with exponentially increasing requests to process DNA samples. This is in part due to the growing importance of DNA evidence in casework since the 1980s but also due to increased demands associated with DNA collection and processing from arrestees and convicted offenders for inclusion in DNA databases. The observations of the research team, along with reports from lab leadership and staff, confirmed that each of

[81] One at the outset of the program, and one after making partial progress on several grant activities.

the sites were faced with intense resource demands associated with their day-to-day workloads upon which was added the implementation of a new program, along with the associated demands of the external evaluation.

Another important issue related to key project personnel. Nearly all sites encountered turnover or a sizable temporary absence of project leaders during the grant period. At one site, this personnel change had dramatic effects on implementation, as nearly all institutional knowledge of the project was lost with that person's departure. While other sites experienced leaves of absence, they were not as disruptive because project information was shared more widely with other lab staff or they occurred later in the implementation process. These sites consequently experienced fewer implementation delays. Further, a wider staff engagement in projects was viewed as a particular facilitator for one site, while the lack of such buy-in was cited as a major detriment for another. Larger staff involvement, at both the analyst and top management level, may be a key strength of such types of laboratory initiatives to ensure greater success and avoid a project crisis if key personnel leave the project.

In addition to internal factors, several of the labs encountered implementation constraints associated with other policies and procedures beyond their control. Public crime labs are not completely independent entities and therefore must adhere to other municipal requirements that can result in delays. The most obvious ones encountered by the sites in this study were local procurement rules. Despite being grant funded, sites had to comply with demanding and cumbersome local request-for-proposal (RFP) and review regulations in order to procure consultant services, equipment, materials and technologies. These added to the delays in implementing components of their efficiency plans. One notable example of this was San Francisco. It took nearly a full year to produce their LIMS development RFP, which was ultimately never completed. Other labs mentioned the procurement process as a challenge to timely implementation. In addition, two labs experienced a lab move or expansion, and one lab faced the demands of external accreditation during the course of the evaluation.

Finally, given the innovative and novel nature of the proposed efficiency approaches, a substantial amount of time and effort had to be devoted to the validation of new instrumentation, software, or robotics to insure equipment sensitivity, reproducibility, and reliability along with compliance with accrediting body standards. Staff also had to be trained on the new processes and processing solutions, which was time intensive, especially when that training was provided in house by senior laboratory staff.

Given these and other factors, the original grant periods of performance had to be extended through multiple no-cost extensions. Understanding these constraints on implementation, NIJ may want to reassess grant award periods, and labs should carefully consider how long it will realistically take to bring project plans to fruition when planning

efficiency approaches in the future, either through this grant program or other funding streams that seek forensic laboratory improvement.

8.1.2 Project Management

Management approaches to the implementation of efficiency improvements through the NIJ grants varied considerably across the sites, particularly in terms of strategic planning and extent of collaboration. Louisiana, for instance, took a systematic approach to designing its intervention, approaching implementation with the idea that they first needed to identify the nature of its processing bottleneck problems. To do so, they adopted a data-driven methodology in which they undertook to empirically describe and measure their bottlenecks in order to craft solutions. Staff perceptions of where process bottlenecks existed were a starting point, but problem identification supported by data and analysis guided solution development. While all labs developed admirable plans for change, not all participated in such an extensive planning process. Consequently, some other projects were deemed ill-suited for incorporation into the current lab's functioning after development and implementation had already begun.

Moreover, labs had varying degrees of collaboration. Louisiana and Kansas City encouraged collaboration for the project activities across multiple organizational levels throughout the lab (Louisiana even involved a group of external stakeholders). Perhaps not surprisingly, both of these labs also had the highest success in terms of implementation. At the other sites, project involvement was less collaborative. For example, at Allegheny County, project management was the primary responsibility of a single manager, with assistance from one of his analysts. When they both left in the middle of the grant period, progress halted until others could learn enough about the program to move forward again. UNT also encountered implementation delays, largely due to a lack of collaboration and buy-in between the research and case-working divisions of the lab. While both of these labs have since worked hard to engage more staff in the grant project, they experienced significant implementation delays that could not be fully overcome in the project period.

This suggests that crime labs thinking about implementing efficiency gains in the future should consider investing in careful, data-driven problem identification in advance of selecting particular solutions for their own unique situations. It also suggests that developing buy-in via collaboration and communication throughout the lab can lead to a smoother and more comprehensive implementation approach and one in which the entire lab becomes invested.

8.2 DNA Processing Outcomes

Most of the sites demonstrated considerable variation in productivity outcome measures across both cases and time. This not only created challenges in detecting patterns, but also

spoke to the complex nature of DNA processing work. Although a seemingly linear process, in actuality, DNA evidence analysis is much more complex and can vary drastically from one case to the next.

Overall, the NIJ grant program appeared to have been associated with mixed outcomes in terms of throughput and turnaround time. Two sites withdrew from the program, and two additional sites were unable to implement substantial grant components within the study period (unsurprisingly, these sites did not show substantial improvements). The research team was able to assess outcomes for the two remaining sites more thoroughly.

Kansas City was able to implement its program components quickly and successfully although its results were mixed. Throughput appeared to slightly increase after the first implementation milestone; however this change was not significant until analyses examined efficiency indices that controlled for budget. While regression analyses provided support for beneficial effects of various components of the grant program, these improvements were often lost in the overall turnaround time or balanced out by other intervention components related to *increased* turnaround times.

In Louisiana, there was evidence of significant increases in throughput and somewhat more modest improvements in turnaround time. The analysis of efficiency indices also revealed positive outcomes. Both Kansas City and Louisiana also had additional project components implemented since the conclusion of the evaluation, and these may contribute to additional improvements in productivity (although they could not be assessed within the current study).

It should be noted that while two of the other sites, Allegheny County and UNT, did not implement the majority of their grant activities into routine casework in time for outcome assessments, it is still possible that these approaches could result in beneficial impacts once implemented. For instance, a small timed experiment performed by UNT found that the newly developed workflow lasted fewer hours than the traditional workflow, particularly during the data interpretation stage. However, until more of the grant outputs are implemented into casework at both of these sites, no conclusion can be drawn on actual outcomes on throughput and turnaround time. It is also possible that other benefits, such as greater accuracy, could have resulted from the grant but could not be directly observed from the study's analysis.

8.3 LIMS Data Limitations

The primary data source for the external evaluation was each laboratory's LIMS. These data required extensive cleaning, recoding, and the creation of new variables before analysis. While this is usually expected to a certain extent for most data sets, the structure and content

of each lab's LIMS revealed limitations of these data for evaluation and performance measurement purposes.

The extracted LIMS data showed considerable variation across sites. Labs varied in the unit of analysis (i.e., tracking by samples, cases, or routings) and level of detail of processing data. The ability to detect change is limited (e.g., turnaround time improvements can be masked) by recording case-level instead of sample-level information, or by reporting stage completion dates without time information. Several sites also reported that key processing fields like "start date" and "end date" were inconsistently used by analysts; further, important information about reruns or whether a case is canceled or put on hold is often not included (or inconsistently tracked).

Much of the information that would be useful for understanding how cases flow through the laboratory (and detecting changes in these workflows) are held in hard-copy case files rather than laboratory LIMS. While this may work well for analysts who can turn to these files for individual case testimony, it does not facilitate easy analysis of data that might be of interest to lab management. If data-entry practices can be improved and LIMS can be developed with some of these considerations in mind (e.g., incorporate more flexibility to report turnaround time with different start/end points, ability to track data by sample *or* case), laboratory managers may find that LIMS can be a useful performance monitoring or management tool. Managers would then have a greater ability to assess internally how organizational changes are influencing laboratory throughput, turnaround time, and backlog.

LIMS users also may benefit from improved linkages between the laboratory and submitting agencies. If the LIMS systems were linked to case outcome data, labs would be able to remove evidence from queues when cases are resolved before evidence testing. For example, Louisiana contacted submitting agencies directly to gather this information. The site reported that this effort allowed them to remove 500 cases from their backlog, without testing, because they had been legally resolved or closed. If this type of case outcome data were entered into the LIMS systems on a continual basis, more evidence submitted but not yet tested could be removed from the testing queue in a more efficient manner.

8.4 Summary

The findings of the Evaluation of the NIJ Forensic DNA Unit Efficiency Program suggest that there is some evidence in support of the hypothesis that crime lab DNA processing can be improved in novel and innovative ways besides simply increasing capacity. However, it should be emphasized that the nature of the research design employed in this particular evaluation precludes making cause-and-effect attributions of the activities funded to the quantitative outcomes observed, because pre/post comparisons and regression analyses cannot completely rule out alternative explanations. Moreover, as was described in some detail under the methodology section of this report, numerous cautions are also in order given

the modest follow-up periods and concerns about the scope, reliability, and validity of the processing data maintained by operational crime laboratories.

Important contributions to the field have occurred through NIJ's grant program. Some prominent examples from each of the sites include drawing attention to the need for more comprehensive and standardized LIMS databases as San Francisco did, widening the eligibility for expert systems use as in Kansas City, UNT's creation of an expert system to decrease time spent on one of the most time-consuming steps of mitochondrial DNA analysis, instituting robotics that allows staff to spend time on other tasks such as at Allegheny County and the other sites, and demonstrating how organizational changes can have significant impacts on DNA processing like in Louisiana. All sites, with the exception of San Francisco's LIMS project, attempted to develop ambitious strategies with entirely new workflows. While such projects are sure to encounter stumbling blocks along the way (as they did), it is possible that only similarly large-scale and paradigm-shifting changes in DNA processing will succeed in matching the growing demands for DNA processing.

More evaluation research is clearly necessary to explore in detail some of the promising findings of this investigation, both from social science and physical science perspectives. This appears particularly warranted given the recent integration of efficiency approaches as part of NIJ's DNA Backlog Reduction Program and the importance of understanding the interrelationships between capacity and efficiency in order to further inform policymakers and practitioners on how to best address the growing backlog and internal crime lab DNA processing issues in the future. As the criminal justice system continues to rely more heavily on the forensic sciences, laboratories will need to pursue new solutions to growing organizational demands. Understanding which of these solutions is most effective is important for both the forensic science and criminal justice fields, as well as the community at large.

APPENDIX A: REFERENCES

Dale, W., Greenspan, O., and Orokos, D. 2006. *DNA Forensics: Expanding Uses and Information Sharing*. Sacramento, CA: SEARCH.

Durose, M. 2008. *Census of Publicly Funded Forensic Crime Laboratories, 2005*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.

Heyne, P. n.d. "Efficiency." *The Concise Encyclopedia of Economics*, www.econlib.org, accessed August 20, 2008.

Hilbe, J. M. 2011 *Negative Binomial Regression.* New York: Cambridge University Press.

Hochendoner, S. June 2011. *Streamlining the DNA Process through Implementation of Automation and Information Technologies.* Document No. 234634.Washington, DC: NCJRS

Horsman, K. M., Bienvenue, J. M., Bienvenue, J. M., Blasier, K. R., and Landers, J. P. 2007. "Forensic DNA Analysis on Microfluidic Devices: A Review." *Journal of Forensic Sciences* 52(4): 784–99.

Hummel, S., and Kennedy, J. 2011. *Increasing Efficiency through Restructuring the Processing of Known Standards.* Document No. 236157. Washington, DC: NCJRS.

McKnight, P. E., and Najab, J. 2010. "Mann-Whitney U Test." *Corsini Encyclopedia of Psychology*. 1.

NCJRS. July 11, 2008. "Search Questions and Answers: Status of DNA Funding." www.ncjrs.gov, accessed August 21, 2008.

National Institute of Justice. 2008. *Increasing Efficiency in Crime Laboratories*. Washington, DC: U.S. Department of Justice.

National Institute of Justice. 2003. *Report to the Attorney General on Delays in Forensic DNA Analysis*. Washington, DC: U.S. Department of Justice.

National Institute of Justice. 2008. *Solicitation: Forensic DNA Backlog Reduction Program.* Washington, DC: U.S. Department of Justice.

National Institute of Justice. 2008. *Solicitation: Forensic DNA Unit Efficiency Improvement*. Washington, DC: U.S. Department of Justice.

Nelson, M. 2010. “Making Sense of DNA Backlogs–Myths vs. Reality.” *National Institute of Justice Journal*, Issue No. 266.

Office of Justice Programs. August 15, 2008. *Request for Quote 2008Q_062, National Institute of Justice – Evaluation of DNA Unit Efficiency Improvement Program.* Washington, DC: Office of Justice Programs, U.S. Department of Justice.

Osborne, Jason W., and Amy Overbay. 2004. “The Power of Outliers (and Why Researchers Should Always Check for Them).” *Practical Assessment, Research & Evaluation* 9(6). Retrieved August 3, 2011 from http://PAREonline.net/getvn.asp?v=9&n=6.

Peterson, J. and Hickman, M. 2005. *Census of Publicly Funded Forensic Crime Laboratories, 2002*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.

President’s DNA Initiative. n.d. “Capacity Enhancement Funding Chart.” www.dna.gov, accessed August 25, 2008.

President’s DNA Initiative. n.d. “How You Can Use the Funds.” www.dna.gov, accessed August 25, 2008.

President’s DNA Initiative. n.d. “Forensic DNA Backlog Reduction Program Legislation.” www.dna.gov, accessed August 25, 2008.

President’s DNA Initiative. n.d. “Program Requirements.” www.dna.gov, accessed August 21, 2008.

Richard, M., and Kupferschmid, T. D. July 2011. *Increasing Efficiency of Forensic DNA Casework Using Lean Six Sigma Tools*. Document No. 235190. Washington, DC: NCJRS.

Roby, R. K., Phillips, N. R., Thomas J. L., Sprouse, M., and Curtis, P. 2011. *Development of an Integrated Workflow from Laboratory Processing to Report Generation for mtDNA Haplotype Analysis*. Document No. 234292. Washington, DC: NCJRS.

Roman, J., Reid, S., Reid, J., Chaflin, A., Adams, W., and Knight, C. 2008. *The DNA Field Experiment: Cost-Effectiveness Analysis of the Use of DNA in the Investigation of High-Volume Crimes*. Washington, DC: The Urban Institute.

APPENDIX B: DATA PROCESSING

PRELIMINARY DATA PROCESSING

The research team needed to perform substantial data processing and cleaning to prepare the data for analyses. This preparation consisted of four steps listed below. Many conversations occurred between the UI research team and the sites to better understand the data and troubleshoot issues as they arose, and the research team is thankful for their guidance through this extensive process.

1. Merging and re-structuring of databases within sites
2. Cleaning data fields and creating new variables
3. Identifying and addressing unfinished cases
4. Identifying and addressing outliers and other data problems

Three of the four sites had multiple datasets (due either to multiple LIMS used within the lab, separate tables produced by a single LIMS, separate tracking spreadsheets maintained by individual analysts, or a change in the LIMS during the study period) which required merging to produce one analytic database per site. Within each site, these datasets were merged and sometimes restructured to align units of analysis (e.g., one dataset was based on case while another was based on sample). Ineligible data were excluded to produce one analytic file tracking relevant DNA work[1] begun in 1/1/2007-1/31/2011.[2] Across the study period, if a lab altered its practice of recording information or changed its LIMS and did not maintain the same data fields, the research team could not use this information in measuring the impacts of the grant program. When multiple dates were encountered for a single sample or case but could not clearly be divided into separate case processing streams (see description of Allegheny County and Kansas City below), the research team retained the earliest date for start points (e.g., submission, assignment) and the latest dates for intermediary or end stages. This was done, because (a) multiple entries often did not have complete information for all stages, (b) when intermediary dates were in one file and flanking dates were in a second file, it was unclear which set of intermediary dates matched to which set of flanking dates, (c) taking only one set of dates and not accounting for other work through additional submissions/re-runs would underestimate the amount of work done at the site, and (d) inconsistently tracked information oftentimes meant that it was impossible to determine the number of re-runs or submissions for a case. This solution resulted in the greatest likelihood of having dates that were not chronologically out-of-order or largely incomplete.

In addition to merging, other data processing occurred to remove any isolated identifying information, categorize string or text variables, clean date fields of trailing and leading text strings, and create new variables from existing data, such as throughput and turnaround time

[1] Examples of irrelevant DNA work include: (a) DNA cases worked outside of the study's reference period, (b) blank, control, or allelic ladder samples, or (c) non-family reference samples for UNT.
[2] Allegheny County and UNT used 2/28/2011 as the end date, because their data was received later.

outcomes, flag and filter variables to identify various types of data, and other key variables. Once the merged databases were finalized, the research team then addressed other issues existing in the databases, including incomplete cases, outliers, and other data problems which threatened the validity of the study's analyses. More details on this data processing work are described below.

Identify and Address Unfinished Cases

One analytic issue confronted by the UI research team was the presence of unfinished cases in the data. These cases could not be simply removed from analyses, because it was more likely that unfinished cases would be concentrated at the end of the study period. *Completed* cases towards the end of the post-intervention period would, of necessity, be cases with shorter turnaround times. Cases taking a longer amount of time would consequently be incomplete at the time of data acquisition. Therefore, removing the unfinished cases would skew the post-intervention turnaround times towards smaller numbers. Such a skew could lead to observing positive intervention effects, when there were, in fact, none. This risk was highest at UNT which had the largest amount (3.7 percent) of unfinished work (see Table 16).

In order to prevent this fallacy, the research team performed median imputation on turnaround times for unfinished cases. A case was defined as "unfinished" if it had started DNA work[3] (evident by the presence of at least one processing stage) but had not yet produced a report or undergone review according to the data. Canceled cases were not included if this information could be parsed out.[4] For these unfinished cases, the team first replaced any missing stage-level turnaround times with the median[5] of all turnaround times for that specific stage in the existing data. These imputed stage-level turnaround times were then added to any existing stage-level turnaround times to produce the overall turnaround time for the sample or case. Utilizing median imputation is a slightly conservative approach in that later, missing stages are estimated based on *all* data, including the pre-intervention turnaround times. However, it still allows for the inclusion of true turnaround times from earlier stages that are not missing and reduces the likelihood of skewing the post-intervention period. Overall, only a very small number of samples/cases/routings (<.01-3.7 percent) were eligible for the median imputation.

Identifying and Addressing Outliers and Other Data Problems

The UI research team conducted a series of diagnostic tests of the data at each site (see Table 16). These exercises revealed some substantial data issues that had the potential to impact the study's analyses. The first issue was the presence of large outliers. Although the number of

[3] No serology cases were incomplete for Allegheny County.
[4] Allegheny County and UNT did not have any information in the LIMS data about case status.
[5] The median was used instead of the mean, because of large outliers present in the data. Details on these outliers and how they were addressed are provided below.

outliers was modest (1.1-5.3 percent), their magnitude was substantial. Standard deviations for each site's estimate of raw turnaround times ranged from 68.81-314.75 days (see Tables 3, 5, 8, and 11), and z-scores of individual turnaround times (overall and stage-level) showed that some turnaround times were as large as 90 standard deviations away the mean (see Table 16). In addition, the minimum and maximum values in Tables 3, 5, 8, and 11 provide additional evidence of the extent of outliers in the data. The largest outliers were in Kansas City and UNT. While it is possible that many of these outliers are, in fact, valid turnaround times, they are likely not business-as-usual for the labs and would serve as a poor comparison against other cases. Furthermore, they could strongly influence the analyses, decreasing the ability to detect true changes across time. [6]

A second identified problem was negative turnaround times. Obviously, it is not possible to spend a negative amount of time on a task; however, two of the sites had a significant number (9.1-15.3 percent) of negative turnaround times. Upon inspection of these negative turnaround times, they were due to dates that were chronologically out of order.[7] While some of these out-of-order dates appeared to be data entry mistakes, delayed entry of dates into the system, or confusion of chronology due to non-linear nature of DNA processing (see *Section 2.3.1, Outcome Data, Measurement Definitions and Complications*), the larger number of negative turnaround times at two of the sites involved systematic issue involving case-based dates. In Allegheny County, cases with multiple submissions often had start/end dates which could not be chronologically linked to intermediary dates. Kansas City had many dates at the sample-level; however the interpretation, technical review, and report stages had both sample-based and case-based dates for the data prior to the new LIMS. When there was missing information for any of these three dates, the case-based date for this stage was included. The majority of negative turnaround time estimates occurred for the turnaround time between the Interpretation and Technical Review stages. More interpretation dates were missing than either the technical review or report dates, resulting in situations where sample-based dates were used for all stages except for the interpretation stages. Many of the negative turnaround times at Kansas City occurred in this circumstance when the technical review date was sample-specific, but the date of interpretation was for the entire case and may have occurred at a later date if there were multiple future submissions.

A final issue investigated was whether any dates were outside of the possible date range. Although the current cases had already been selected based on whether their start dates were in the study reference period, no selection was done based on other dates. Specifically, dates were flagged as being out-of-range if they were past June 30, 2011. Only two sites had any

[6] Outliers can negatively impact analyses by violating the normality assumption, providing more weight to outlier values due to the squaring of residuals, and reducing power. An informed removal of outliers can lead to more accurate parameter estimates (Osborne & Overbay, 2004).

[7] The percentage of negative turnaround times does not exactly match to the percentage of out-of-order dates, because (a) some dates were not used to calculate turnaround times and (b) negative overall turnaround times would not be identified as a sample/case with out-of-order dates because this determination was made based on adjacent dates.

samples/routings where this occurred. These identified situations were obvious data entry errors with years beyond 2011.

The research team treated all of the above scenarios with a similar cleaning protocol. First of all, if any one case/sample/routing had three or more of each type of problem, it was excluded from all turnaround time analyses under the thought that the data was too "messy" to draw conclusions from and it would be unclear which dates were accurate. If a case/sample/routing had two or more outliers *and* two or more negative turnaround times, it was also excluded. Individual data problems (i.e., an outlier, negative turnaround time, chronologically out-of-order date, or date which was outside of the possible date range) were replaced with a missing value, although the remaining values for the case/sample/routing were maintained. All cases undergoing median imputation, exclusion from analyses due to multiple and substantial data quality problems, or individual value removal were flagged as having altered data. Original data were preserved.

Table 16. Frequency of Data Issues by Site

Site	Sample Size	Unfinished Cases	TAT Outliers	TAT z-score range	Out-of-order dates	Negative TAT	Out-of-range dates
PA	All Sero/DNA forensic evidence *(N=1518) (N<400 for cases with DNA)*	0.3%	4.7%	(-16.9, 17.5)	9.0%	9.1%	0%
KS City	All DNA *(N=10,296)*	0.1%	5.1%	(-92.1, 91.4)	14.3%	15.3%	<0.1%
UNT	All Fam Ref Samples *(Rout=3429 Case=3079)*	3.7%	1.1%	(-51.6, 54.2)	0.5%	0.5%	0.1%
LA	All DNA forensic evidence *(N=4325)*	<.01%	5.3%	(-1.5, 23.4)	0.1%	<0.1%	0%