Developing Criteria for Model External DNA Proficiency Testing

Final Report

National Institute of Justice
Grant No. 96-DN-VX-0001

Joseph L. Peterson, D.Crim.

R. E. Gaensslen, Ph.D.

with the research assistance of

George Lin, M.S.
Monica Ho, M.S.
Ying-yu Chen, M.S.

University of Illinois at Chicago

May 2001

TABLE OF CONTENTS

TABLES AND FIGURES

APPENDICES

EXECUTIVE SUMMARY

The Blind DNA Proficiency Testing Feasibility study, funded by the National Institute of Justice (NIJ # 96-DN-VX0001), had six primary objectives:

1. Appoint and utilize a National Forensic DNA Review Panel (NFDRP) (See Appendix B for names and affiliations);

2. Review the blind and open proficiency testing literature in the clinical and forensic fields;

3. Survey forensic DNA testing laboratories, selected law enforcement agencies, defense attorneys and independent expert witnesses;

4. Explore alternative strategies for a potential blind proficiency testing program for forensic DNA testing laboratories in the United States;

5. Conduct limited blind proficiency tests in a selected sample of forensic DNA testing laboratories to assist in determining the overall feasibility of a national program; and

6. Present findings, options and recommendations to the NFDRP and to the Director of NIJ.

The NFDRP consisted of twenty members, including most of the members of the DNA Advisory Board at the time this project was initiated, as well as six other persons representing interested groups and organizations. The panel was chaired by Dr. Joshua Lederberg. Four meetings of the NFDRP were held at which progress on the project was reported and discussed. Operationally, the project was divided into two principal Phases (1 and 2). Phase 1 involved an initial survey of forensic DNA testing laboratories and law enforcement agencies that submit evidence to them. The literature review was begun in Phase 1 and completed in Phase 2. The initial round of ten proficiency tests was completed in Phase 1 and preliminary findings drafted. An initial set of recommendations was formulated and communicated to NIJ in a major Phase 1 final report.

With continued funding, the project team was able to conduct additional surveys of DNA laboratories, and was also able to survey defense attorneys and independent experts, plus review various random reanalysis schemes. An additional round of five, more challenging, blind tests was then run and completed. In total, 15 blind tests were satisfactorily issued and results received. Replicate blind "cases" were likewise issued to 7 reference laboratories, and results received. This executive summary reviews the major findings of the research, addressed in detail in the body of the report.

Phase 1

Creation of the NFDRP was a key step in the project, since it represented the various constituencies who would perforce be involved in any blind DNA proficiency testing program.

The panel membership also supplied broad perspectives on proficiency testing issues and represented the various knowledge, skills and logistical areas that would contribute to the overall design of a workable proficiency test program.

The literature review summarizes the history of proficiency testing efforts, identifying significant regulatory and scientific landmarks that have placed proficiency testing as the centerpiece of most quality assurance (QA) programs within clinical, medical, urine drug testing and forensic areas. Attention is devoted to QA, standards setting, and proficiency testing experiences in forensic DNA testing over the past decade. Given the specific objectives of this project, particular attention is paid to blind testing in related fields. Because of the importance and relevance of efforts in the past fifty years to regulate the clinical laboratory field, the review begins with an overview of significant achievements in this area.

In the nationwide survey of forensic DNA laboratories in Phase 1, one goal was to determine the procedures laboratories employ in the routine handling, processing and reporting of evidence. This knowledge was needed to determine if and how blind samples might be introduced into forensic laboratory operations. Another goal of the survey was to determine the extent to which laboratories were involved in QA and proficiency testing. The objective of the law enforcement agency survey was to determine the willingness of these agencies to participate in the study and, if not, reasons for their refusal/reluctance.

Of 39 potential target DNA testing laboratories that agreed to participate in the actual testing in Phase 1, ten were selected for participation as testing targets. Project staff also worked with the contractor tasked with providing blind tests to the FBI Laboratory. Altogether, ten proficiency tests were submitted to DNA labs. Of those ten, one "case" was recognized by a target laboratory as a blind test, and two of the cases were not completed (one because of backlog, and another because of a communications problem between the project team and the law enforcement agency). Two of the remaining 39 labs were used as reference labs in Phase 1. Blind proficiency testing is discussed in the report in terms of the purposes of proficiency testing and its role in quality assurance, and in comparison with declared proficiency testing. The issues, problems and logistics of a national blind testing program under several models and assumptions are discussed, along with the data that might be obtained from such a program. The estimated costs of blind proficiency testing, on a cost-per-test basis, is estimated to vary between 7 to 50 times the cost of declared proficiency tests, depending on various assumptions made in computing the estimate and on the complexity of the blind testing undertaken. Annual costs for administering a single blind test to each of an estimated 150 DNA testing laboratories nationally are projected to be from $ 220,000 to $ 1,510,000, and from $ 450,000 to $ 3,020,000 per year if two blind tests per year were to be administered.

At the final meeting of the NFDRP in Phase 1, the advisory group considered the draft Phase 1 Progress Report, and engaged in a daylong discussion of the issues, complexities and problems connected with a large-scale program of blind proficiency testing. The panel's recommendations to the National Institute of Justice (and to the DNA Advisory Board) were formulated, in order that NIJ could prepare its recommendations and report to Congress (to comply with the mandate of the DNA Act of 1994).

Major points of discussion surrounding the formulation of the final recommendations are presented below to help provide some insight into their logic and language.

Some NFDRP members suggested that there is little evidence that there is a serious problem in DNA testing laboratories (for which blind proficiency testing is proposed as the solution). A regularly scheduled blind proficiency-testing program could draw resources away from existing QA/QC procedures and practices - thus possibly not yielding any net improvement. Other panel members believed quite firmly that the information to be derived from a large-scale blind proficiency-testing program that is not available in other ways, or through declared testing, is essential to properly monitor the laboratories. The principal advantages of blind testing are seen as: (1) Superior to open tests for detecting errors because analysts are unaware they are being tested; (2) Tests "the whole system" by allowing a more expansive review of the handling, selection of evidence for analysis, analysis itself, and interpretation of results, than does open testing which is restricted to testing analytical procedures; (3) Creates heightened vigilance on the part of examiners since they never know which case might be a test; and (4) Builds greater public confidence in the forensic DNA testing process.

The strongest reservations against implementing a broad program of blind proficiency testing using law enforcement or conduit laboratory models are its anticipated cost, complex logistics, and potential CODIS problems. Another weakness of blind testing is that it does not address the initial crime scene and evidence collection steps taken by the police which, in fact, may be the weakest link in the entire process. There are additional problems, including the difficulties in creating replicate challenging cases, donor and data basing issues involving fictional cases, and the ethical implications of a process that requires law enforcement personnel to deceive laboratories.

Many panel members were neither strongly opposed nor strongly in favor of a large-scale program. To the extent that a program was to be recommended, most members gravitated toward the simplest, least costly method for achieving the stated goals of a national blind proficiency-testing program.

Because of the complexities, costs and problems associated with the introduction of blind proficiency tests via law enforcement agencies (Blind / LE), conduit laboratories (Blind / CL), and/or Blind Analyst modalities, the focus turned to consideration of the "random reanalysis" model. Interpreted broadly, this model could be simply termed "reanalysis," and would include complete technical audits of previously completed cases, with retyping evidence an option if it were seen to be necessary or important by the auditor. The reanalysis could include an evaluation of the judgments used in selecting specimens for typing from the original evidence, provided the original evidence were still available. Thus, in many respects, reanalysis could provide virtually the same information that would be obtained by the more complex blind testing models, and is much simpler to administer. The costs of reanalysis could be reduced significantly, if retyping of evidentiary items was not required in every instance of reanalysis, but only when an auditor felt that there was cause to do so.

Under the reanalysis model, cases for audit could be selected on some basis (e.g., an interested party believed there to be a problem with the case) or could simply be a random sample. For this model to fall under the definition of "external," parties external to the laboratory/prosecution team would have to be involved in the selection of the case, its review, and its reanalysis. Guidelines would have to be established concerning potential auditors, and case selection criteria. A randomly sampled case or cases is seen as the better approach (given the difficulties deciding who would select or on what basis a case would be selected), with the one limiting condition that the evidence must still be available in order to do the complete audit /

reanalysis. Guidelines may be needed, therefore, addressing the retention of samples for possible re-analysis, in addition to present practices. This blind-testing model does not, therefore, adequately cover cases in which the quantity of biological evidence is very limited. To avoid legal problems that might result if there were a discrepancy or discordance between the original casework and the reanalysis results, cases for reanalysis should probably be chosen from a pool of adjudicated cases.

It is not immediately clear who could perform this audit / reanalysis. It might logically be made a part of a currently required DNA section administrative audit by ASCLD-LAB in accredited laboratories, but ASCLD-LAB may not have the resources to add this function given its already stretched resources. In addition, because a technical review and possible reanalysis is being proposed, the reviewers must have the appropriate scientific qualifications. For credibility reasons, independent DNA analysts should be involved in the review.

At the present time it is not known what percentage of cases should be sampled for reanalysis/audit. An arbitrary figure could be selected, e.g., 5-10 cases per analyst per year, a percentage of cases worked by analysts, or some figure approximating the percentage of cases independent experts believe are problematic.

There was discussion concerning how the "feasibility" of blind testing should be interpreted, i.e., is it synonymous with "possible," or does it imply "practicable in terms of costs and logistics" in addition to "possible." The panel adopted the latter definition: possible and practicable in terms of costs and logistics. Under that definition, a national blind DNA laboratory proficiency testing program employing what we have called the Blind / LE, Blind / CL, and Blind Analyst models is not feasible at this time.

The recommendations that grew out of the above discussions and submitted at the close of Phase 1 of the project were as follows:

1. The accreditation system and associated quality assurance guidelines of the DNA Advisory Board need to be given the opportunity to take hold.

2. It is recommended that the DNA Advisory Board generate guidelines for more stringent external case audits for use by ASCLD-LAB, or another relevant accrediting body, as part of the accreditation process. The external case audits should be conducted regularly and serve as a measure of how well accreditation and its associated requirements are working in a quality assurance context.

3. In the extreme, blind proficiency testing is possible, but fraught with problems (including costs), and it is recommended that a blind proficiency testing program be deferred for now until it is more clear how well implementation of the first two recommendations are serving the same purposes as blind proficiency testing.

Phase 2

Phase 2 of the blind proficiency testing feasibility project included a number of tasks designed to help answer some of the questions raised in the advisory panel's discussions, and thus help to guide future policy makers on questions about a national blind testing program. These activities included a round of additional blind testing, a closer consideration of the re-analysis program, an estimate of the fraction of worked DNA cases that are reviewed/reanalyzed, and the extent to which original evidence items are still available in worked DNA cases that have been adjudicated.

In Phase 2, additional blind trial proficiency tests were designed and submitted to forensic science laboratories that indicated a willingness to take part. The objective of these Phase 2 tests was to gather preliminary data on the feasibility of the accurate replicate-manufacturing of materials in cases that were more complicated than those used in Phase 1. We decided to develop case scenarios and manufacture Phase 2 blind tests around a single evidence item on which had been deposited two persons' blood, that of the "victim" and that of the "suspect." In every case, the scenario involved assault, attempted sexual assault and/or home invasion, and sharp-force injuries (inflicted by a knife that was not recovered) to both parties. Some resulting bloodstains on the pants of the "victim" were from the "victim" while others originated from the "suspect." The "cases" were somewhat more challenging from a criminalistics point of view; that is, labs would have to make decisions about which stains to examine. Five such blind "cases" were constructed and submitted to forensic DNA laboratories through law enforcement agencies. Because part of the challenge in this phase was the reproducible replicate manufacturing of evidence with bloodstains that had to exhibit a pattern consistent with the case scenario, we chose to manufacture a total of ten case items. Five were used in the blind test "cases" submitted to forensic labs, and five were submitted to reference laboratories. All the labs participating in Phase 2 had agreed in advance to be potential test candidates (or reference laboratories), just as was the case in Phase 1. The Phase 2 tests were manufactured and administered in a single cycle. Consequently, in Phase 2, we focused attention on whether evidence items representative of more challenging cases could be replicate manufactured with sufficient reliability to insure uniform results from competent laboratories. This task was accomplished. At least insofar as relatively uncomplicated sets of bloodstain patterns on items are concerned, replicate evidence manufacturing is possible, although it is labor intensive.

Also, in Phase 2, surveys of forensic DNA laboratories focused on their sample retention practices, and internal and external reviews of casework. Evidence retention is a necessity if that evidence is to be re-examined as a quality assurance mechanism. Our survey indicated that more than 90% of laboratories retain the original specimen, extracted DNA, or both, and the great majority of labs reported only about an average of 5% of their DNA analyses involved total consumption of biological samples of interest. Almost all of the responding labs estimated that between 95% and 100% of their DNA cases were subjected to internal reviews, while a much lower percentage receive external audits. Rates of re-testing of evidence for both internal and external audits are quite low, with most of these re-tests agreeing with the original reports.

Six defense attorneys and 11 expert witnesses with experience in reviewing cases involving DNA evidence also returned surveys in Phase 2. These individuals, unlike the DNA testing laboratories surveyed in Phase 1 and 2, reported they often detected problems with the original testing performed in the government laboratories. Half of the responding defense attorneys

indicated that they found discrepancies between re-test results and original reports. Many found fault with DNA laboratory procedures, many questioned the validity of lab reports, and one pointed out the need for defense attorneys to have greater opportunities for assistance from technical experts.

Independent expert witnesses reported rates of re-testing of cases to be very low. Only one-third of the expert witnesses reported performing re-tests on DNA cases, and of these, they only performed re-tests in about half of the cases they reviewed. They had trouble about gaining access to the primary testing labs' documentation that they attributed to laboratories' fear of criticism from outside experts. They also reported cases where critical evidence had been totally consumed, but that this problem had lessened with the growing use of PCR and laboratory protocols requiring labs to save portions of the evidence to allow for re-analysis.

There is little disagreement that a proficiency testing program is a critical quality assurance tool for analytical laboratories and for forensic science laboratories in particular. This project focused on the feasibility of blind proficiency testing and on the advantages and disadvantages of blind vs. open proficiency testing. We have shown that blind proficiency testing is possible, although it is not successful in every trial. Compared with open testing, it is more complex and significantly more expensive. It also tests more components of a laboratory system, whereas open proficiency testing primarily tests the accuracy of analytical results. We have also shown that moderately more complex evidence requiring more than perfunctory judgment by analysts can be replicate manufactured successfully. A random audit / reanalysis program of blind proficiency testing is less complicated than a program where case items are manufactured and submitted through normal law enforcement channels, and may provide almost the same information from a QA viewpoint. Such a program is nevertheless labor intensive and significantly more expensive than open proficiency testing. The limited number of tests done in the project yielded uniformly accurate results and interpretations, and one purposely constructed cross-state case to case CODIS match was found by the CODIS system.

I: INTRODUCTION

A. Introduction

Forensic DNA typing has emerged as one of society's most powerful tools to determine a suspected offender's involvement in a criminal act. DNA technology was first introduced in a United States criminal court case in 1986 and, since then, has become widely accepted by the law enforcement and the judicial systems as a means for positively identifying individuals (and excluding) suspected of committing crimes and exonerating persons convicted of crimes largely on the basis of faulty testimony. Its introduction constitutes one of forensic sciences' greatest evidentiary breakthroughs in this century.

Some of the advantages of DNA analysis over traditional serology are: (1) it is highly informative in associating persons with their blood and other biological fluids; (2) DNA is identical in all cells allowing it to be used in testing a broad range of biological trace evidence; (3) DNA is more stable than enzymes and proteins allowing it to be used with highly degraded samples; and (4) it is extremely sensitive and able to be used with very small samples. Beginning in the mid-1980s, the legal system raised important questions regarding the validity, reliability, and quality assurance of DNA analysis.

B. Validity of Forensic DNA Testing

Several study commissions have evaluated the forensic DNA technology. The Office of Technology Assessment (OTA) of the U.S. Congress issued a report in 1990 which strongly endorsed forensic DNA testing declaring that, "molecular and genetic principles underlying DNA techniques are solid and can be successfully applied to casework. Forensic uses of DNA tests are valid (OTA, 1990)." The National Research Council (NRC) issued its report in 1992 entitled DNA Technology in Forensic Science (NRC, 1992) which agreed with the OTA report that forensic uses of DNA tests are valid and reliable when properly performed. Still, these organizations identified several areas of concern regarding the reliability of results and the quality assurance procedures in the laboratory.

C. Reliability and QA of Forensic DNA Testing

The reliability of DNA testing results has been rigorously scrutinized by the courts. In part, this has stimulated laboratories to implement quality assurance procedures. In 1989, in *New York v. Castro*, Judge Sheindlin ruled that inadequate quality assurance steps were sufficient to hold DNA evidence inadmissible (see Case Citations - Quality Assurance). In November of 1989, the Minnesota Supreme Court, deciding *State v. Schwartz* (see Case Citations - Quality Assurance) became the first appellate court to reject DNA results because of the testing laboratory's errors in an early proficiency testing program performed by the California Association of Crime Laboratory Directors and that the laboratory had not used relevant protocols developed by the FBI. Forensic laboratories were on notice that demonstrating reliability was going to be a crucial issue for determining DNA's legal admissibility.

The forensic DNA community responded with QA guidelines, most notably, the Technical Working Group on DNA Analysis Methods' (TWGDAM) "Guidelines for a Quality Assurance Program for DNA Analysis" (1988). These guidelines address the training and qualifications of personnel; documentation; validation of methods and procedures; evidence handling procedures; audits; analytical procedures; and proficiency testing. The section on proficiency testing states, "Participation in a proficiency testing program is a critical element of a successful QA program

and is an essential requirement for any laboratory performing forensic DNA analysis" (TWGDAM, 1995). These voluntary guidelines also have been widely accepted by the courts as standards for forensic DNA testing.

The ASCLD Laboratory Accreditation Board (ASCLD-LAB) requires laboratories seeking accreditation (also voluntary) to establish and maintain a "quality system", appropriate for its casework. Proficiency testing is cited as an "integral component" of QA programs and requires laboratories to subscribe to an external proficiency test provider in all disciplines in which they seek accreditation (American Society of Crime Laboratory Directors, 1994).

The American Board of Criminalistics (ABC) certifies individuals based upon their educational background, experience, and performance on a written examination. Recognizing the importance of proficiency testing, those seeking Fellow status in a specialty area like DNA testing must also supply proficiency test performance results, and must submit acceptable PT results annually thereafter to maintain their certification.

The Federal Judicial Center's 1994 Reference Manual on Scientific Evidence (McKenna et al, 1994) written to help federal judges interpret and manage scientific evidence, includes a chapter on DNA evidence. Among various recommendations, it suggests the court inquire if the pertinent testing laboratory has demonstrated an appropriate record of proficiency and quality control, "to permit confidence that the tests were conducted properly" (McKenna et al., 1994). The second edition of the Center's manual provides a more detailed description of how proficiency testing has been incorporated into existing accreditation, certification, and standards setting programs (Kaye and Sensabaugh, 2000). The NRC in its 1996 report also recommended that "laboratories should participate regularly in proficiency tests, and the results should be available for court proceedings."

Several of these organizations have also stressed the utility of blind proficiency testing. TWGDAM noted it is "highly desirable" for the DNA laboratory to participate in a blind proficiency testing program that "realistically simulates" actual casework in order to evaluate "all aspects of the laboratory examination procedure" (TWGDAM, 1995). With regard to proficiency testing, the NRC endorsed the TWGDAM guidelines for quality assurance, which is to say they recommended regularly scheduled proficiency testing as a way of measuring laboratory error rates (false positives and false negatives). The earlier 1992 NRC report recommended that error rates be "continually estimated in blind proficiency testing" (NRC, 1992). The Committee commented that errors occur in the best laboratories and that "error rates" need continuous review and adjustment: "One purpose of regular proficiency testing under standard case conditions is to evaluate whether and how labs have *taken corrective action to reduce error rates"* (NRC, 1992). They noted such tests would ideally involve blind tests of representative case materials.

Finally, the authors of the Reference Manual on Scientific Evidence cited earlier suggest that courts may want to give more weight to blind test than open test results.

D. The DNA Identification Act of 1994

Mandatory external blind proficiency testing was proposed during a joint hearing on forensic DNA analysis in 1991 before the Subcommittee on Civil and Constitutional Rights of the U.S. Congress. This proposal was based on testimony that clinical laboratories are mandated by CLIA to participate in blind proficiency testing (U.S. Congress, Hearings, No. 30, 1991). (Note: CLIA does not have such a requirement, nor does such a national program of blind testing exist.) Additional testimony was offered including recommendations: to create an independent

board of scientists to set standards; to license forensic laboratories and their personnel; and to address privacy issues surrounding DNA type databasing. About the same time, in the state of New York, a model statute was passed by the legislature to regulate all forensic laboratories in New York State. In response to these proposals, the U.S. Congress passed the DNA Identification Act of 1994 (P.L. 103-322, 1994) as part of the Violent Crime Control and Law Enforcement Act of 1994. The DNA Act established a framework (similar to New York State's) for setting standards on quality assurance and proficiency testing in forensic DNA typing laboratories. The law created the DNA Advisory Board (DAB), specifying that members be appointed by the Director of the FBI. The DAB consists of forensic scientists, molecular and population geneticists, and others knowledgeable in law and ethics. The law states, "The advisory board shall develop, and if appropriate, periodically revise, recommended standards for quality assurance, including standards for testing the proficiency of forensic laboratories, and forensic analysts, in conducting analyses of DNA" (P.L. 103-322, 1994). In addition to standards setting, the DNA Act further authorized the appropriation of grant monies for laboratories to establish and/or improve DNA testing services and DNA databases and repositories. One condition for eligibility for those monies is the implementation of a strict QA program adhering to the DAB's standards for DNA testing. The standards include participation in a regularly-scheduled, external proficiency testing (PT) program.

The DNA Act also directed the National Institute of Justice to establish a national blind proficiency testing program for DNA analyses in public and private forensic science laboratories. Because no such program existed, the NIJ, instead, opted to fund a study of the feasibility of blind proficiency testing and, in 1995, issued a solicitation for grant proposals titled "Developing Criteria for Model External DNA Proficiency Testing." In the solicitation, key points of the proposed research included: reviewing the state of science of blind and open external PT programs; analyzing and comparing external PT programs in terms of reliability and validity of performance; and developing a candidate model for nationwide blind proficiency testing.

E. Research Plan

The primary purpose of this four-year research study, including Phase 1 and Phase 2, was to test the feasibility of national, blind DNA proficiency testing in public and private forensic science laboratories. Six major activities were proposed for this purpose:

1. Creation of a National Forensic DNA Review Panel (NFDRP),
2. Literature review of blind and open proficiency testing in the clinical, forensic, and related laboratory fields,
3. In-depth survey of forensic science laboratories, selected law enforcement agencies, defense attorneys, and expert witnesses,
4. Design alternative blind proficiency testing strategies,
5. Field tests of several, small scale blind PT strategies and documentation of results,
6. Final report and recommendations.

The creation of the NFDRP was an important step as it was composed of representatives who would be involved in any blind DNA proficiency testing program. The panel also represented the various knowledge, skills and logistical areas that would be needed to design a workable proficiency test program. The various individuals on this panel also needed to provide broad perspectives on issues in this project, looking beyond their own jurisdictional and organizational interests and concerns.

The literature review identifies significant regulatory and scientific landmarks that have placed proficiency testing squarely at the core of most quality assurance (QA) programs in the clinical, medical, urine drug testing and forensic areas. Particular attention is paid to QA, standards setting, and proficiency testing experiences in forensic DNA testing over the past ten years. Blind testing is examined in particular given the thrust of this project. The review begins with a brief overview of significant achievements in regulating the clinical laboratory field given its relevance to various quality assistance efforts.

The nationwide survey of forensic DNA laboratories in Phase 1 determined the procedures laboratories employ in the routine handling, processing and reporting of evidence. This knowledge allowed us to explore how blind samples might be introduced into forensic laboratory operations. This survey also determined practices followed by laboratories in terms of QA and proficiency testing. The law enforcement agency survey determined the willingness of police agencies to participate in the study.

In Phase 2, newly designed surveys were distributed to existing and newly established forensic DNA laboratories to obtain additional information from them. The goal of the Phase 2 survey focused on laboratories' sample retention practices, which is a crucial step in determining the possibility of re-tests. In addition to sample retention practices, the frequency, nature, and outcomes of internal and external audits were also examined. Surveys on defense attorneys and expert witnesses explored the nature and percentage of their DNA cases which received additional re-tests and various types of reviews.

Candidate blind proficiency testing schemes were designed with varying degrees of blindness ranging from ones in which no one in the laboratory organization realized they were being tested, to those where persons/supervisors other than the analyst know it's a test. In addition, case types (e.g., sexual assaults, blood transfer cases) and DNA databanking concerns were considered in developing these blind proficiency tests. In Phase 1, eight separate blind proficiency tests were issued. Two additional blind PTs were developed in collaboration with the contractor responsible for submitting blind DNA PTs to the FBI Laboratory. These tests were submitted to both commercial laboratories and public laboratories (federal, state and municipal). In Phase 2, blind proficiency tests were issued to five additional laboratories, and five were submitted to reference laboratories. The results of these limited blind PTs are documented in remaining sections of this final report.

F. Conclusion

The data derived from these activities/tasks assisted the project staff to develop recommendations about the implementation of blind DNA proficiency testing in forensic science laboratories on a nationwide basis. The investigation also yielded information about the costs of the testing, the problems in maintaining "blindness" of the testing, and a basis for designing and administering such tests on a large scale. Because the National Forensic DNA Review Panel helped plan operational details, monitoring progress throughout the project, and in formulating final recommendations, we believe it also enhances the credibility of the findings, and helps insure acceptance of our findings by regulators, legislators, and the forensic science community. The panel and project staff devote particular attention to the issue of error rates in DNA testing -- ways the scientific community defines error in an applied laboratory setting, and the role proficiency testing plays in operationalizing such measures. Findings in the six major activities of our research plan are detailed in the following chapters of the Final Report and Recommendations.

II: LITERATURE REVIEW

A. Introduction

Forensic DNA testing has become one of the most important laboratory techniques employed in the identification of individuals in criminal cases. As such, the forensic community has devoted considerable attention to this area to ensure the quality, integrity, and reliability of the DNA testing results. With the passage of the DNA Identification Act of 1994, the United States Congress added its voice to the call for higher standards in DNA testing by providing funds to state and local governments to improve the capabilities of DNA laboratories, the creation of a DNA Advisory Board by the Director of the FBI to require laboratories' conformance to quality assurance standards, and the investigation of the feasibility of blind proficiency testing

Our research proposal to the National Institute of Justice to investigate the feasibility of blind proficiency testing included an initial literature review. Objectives of this present literature review include summarizing the history of proficiency testing efforts, including significant regulatory and scientific landmarks that have placed proficiency testing as the centerpiece of most quality assurance programs within the clinical, medical, urine drug testing and forensic areas. Attention is also devoted to the quality assurance, standards setting, and proficiency testing experiences in the forensic DNA testing area over the past decade. In addition, the issues of laboratory error rate and proficiency testing as indicators of laboratory performance are addressed. Given the specific objectives of this research project, particular attention is paid to blind testing efforts in related fields with a comparison of the merits of blind and open proficiency testing. Finally, because of the importance and relevance of efforts in the past fifty years to regulate the clinical laboratory field, a section is devoted to reviewing the significant achievements and milestones in this area.

B. Quality Assurance / Quality Control in Forensic DNA Testing Laboratories

Forensic deoxyribonucleic acid (DNA) testing is a powerful technique for the identification of individuals involved in parentage disputes and criminal matters. Since the results of DNA testing can have a profound impact on the lives of litigants, the forensic and parentage DNA communities must assure all parties in these disputes that test results are accurate and reliable. This assurance keys upon the ability of DNA testing laboratories to meet or exceed rigorous quality assurance (QA) and quality control (QC) standards established via consensus among laboratory analysts and overseeing agencies (Anderson et al., 1995). Quality assurance (QA) has been defined as "all planned or systematic actions necessary to provide adequate confidence that a product or service will satisfy given needs" (Freund, 1985). A comprehensive QA program for laboratories performing analytical testing (including forensic DNA testing) would typically include periodic laboratory audits by external specialists, maintenance of current and clearly written protocols, preparation of QA reports, troubleshooting, equipment maintenance and calibration, methodology development, personnel training, continuing education, laboratory safety and quality control (QC) (Kirby, 1990). Quality control (QC) has been defined as the operational techniques and activities necessary to sustain the quality of a product or service and is the aggregate of processes and techniques so derived to detect, reduce, and correct deficiencies in an analytical process (Freund, 1985).

As such, QC involves all aspects of testing which have a direct impact on the accuracy and precision of the final laboratory result. A comprehensive QC program describes the types of controls and standards which are appropriate for each test, the proper documentation of reagents and supplies, and means of verifying that: (a) the correct specimen is being analyzed; (b) the specimen is adequate for that particular test; (c) the specimens are not inadvertently mixed or cross contaminated within the test laboratory; (d) the results are being correctly interpreted; and (e) the written protocols are being followed without deviation (Kirby, 1990).

The first applications of DNA testing to disputed parentage and forensic identification were conducted by private, commercial laboratories. Quality assurance procedures in these laboratories were inspired in part by the anticipation of rigorous scrutiny from the courts, especially in criminal cases. Hence, these laboratories were the first to establish QA/QC programs for forensic DNA testing (Anderson et al., 1995). It is worth noting that parentage testing on one hand, and genetic-marker testing in criminal case evidence on the other have, for the most part, been accomplished in different sets of laboratories in the United States. In the United States, most parentage testing developed in clinical laboratory settings because of the historically close relationship between blood grouping and histocompatibility testing for clinical purposes and parentage testing. Very few of these laboratories test evidence from criminal cases. Similarly, very few public forensic science laboratories that examine criminal case evidence are involved in parentage testing.

The great majority of courts that have since addressed DNA admissibility issues have found that the underlying theory of DNA typing is sound and that the laboratory procedures and techniques are valid, reliable and generally accepted if competently performed (NRC, 1992). There have been some courts, however, that questioned the methods used to compute frequencies of occurrence of DNA types in "match" cases (NY v. Wesley, 1988; U.S. v. Yee, 1991; see Case Citations - Statistics). This led to the creation of another panel of experts by the National Research Council to address such statistical questions. This panel's final report resolved (most) of the judicial concerns (NRC, 1996).

The first published set of quality assurance standards for DNA testing were those of the American Association of Blood Banks (AABB) (AABB, 1991). These standards were created for parentage testing laboratories performing restriction fragment length polymorphism (RFLP) analysis. Recently, the AABB revised the parentage testing standards to include several key provisions for polymerase chain reaction (PCR) based testing. The American Society of Histocompatibility and Immunology (ASHI), a group that has splintered off from the College of American Pathologists (CAP), has established standards for Human Leukocyte Antigen (HLA) DQA locus testing (ASHI, 1995).

The Technical Working Group on DNA Analysis Methods (TWGDAM), created in 1988, provides consensus guidelines for QA/QC within the forensic science DNA typing laboratory community. This group, established as an independent organization and funded by the Federal Bureau of Investigation, is composed of leading forensic scientists involved in forensic DNA testing, and is divided into four subcommittees: RFLP analysis, PCR-based testing, mitochondrial DNA sequencing, and QA/QC. TWGDAM first published its "Guidelines for a Quality Assurance Program for DNA Analysis" in 1991which have undergone several revisions with the current version appearing in April 1995 (TWGDAM, 1995). These guidelines include statements about the training and qualifications of personnel; documentation; validation of methods and procedures; evidence handling procedures; audits; analytical procedures; and

proficiency testing. These guidelines have been widely accepted by the courts as standards for forensic DNA testing. The International Society for Forensic Haemogenetics (ISFH) has also published guidelines for forensic DNA testing (Proceedings of the International Symposium on the Forensic Aspects of DNA Analysis, 1989). In addition, ISFH has a quality assurance commission and a subcommittee, known as the European DNA Analysis Panel, which provides quality oversight for DNA testing in Europe and coordinates the development of new methods (Anderson et al., 1995).

The DNA Identification Act of 1994 (P.L. 103-322, 1994) established a federal framework for setting standards on quality assurance and proficiency testing in forensic DNA typing laboratories. The law created the DNA Advisory Board (DAB) and specifies that members be appointed by the Director of the FBI. The DAB consists of forensic scientists, molecular and population geneticists, and others knowledgeable in law and ethics. The law states, "The advisory board shall develop, and if appropriate, periodically revise, recommended standards for quality assurance, including standards for testing the proficiency of forensic laboratories, and forensic analysts, in conducting analyses of DNA" (P.L. 103-322, 1994). Since its appointment, the DAB has been developing standards for forensic DNA testing. The DNA Act further authorized the appropriation of grant monies for laboratories to establish and/or improve DNA testing services and DNA databases and repositories. One condition for eligibility for those monies is the implementation of a strict quality assurance program adhering to the DAB's standards for DNA testing. The standards include participation in a regularly-scheduled, external proficiency testing (PT) program.

The National Institute of Standards and Technologies (NIST) has also played a role in quality assurance and quality control in forensic DNA typing laboratories. Because NIST has extensive experience in evaluating technologies, and has conducted significant independent research on measurement precision in various DNA typing methodologies, it has developed a program to test and develop a series of standards and controls for use in DNA typing. Its DNA profiling standards for RFLP and PCR are intended for use in "standardization of forensic and paternity QA procedures" and "for instructional law enforcement or nonclinical research purposes" (NIST, "Certificate of Analysis," 1992).

In 1998, the National Institute of Justice created the National Commission on the Future of DNA Evidence. The group's objective is to develop policies that will maximize the value of DNA in the criminal justice system. Among the issues to be addressed by this commission will be five specific areas: (1) the use of DNA in post-conviction relief issues, (2) legal concerns including *Daubert* challenges and the scope of discovery in DNA cases, (3) criteria for training and technical assistance for professionals involved in crime scene/evidence gathering procedures, (4) essential laboratory capabilities in the face of emerging technologies, and (5) the impact of future technological developments on the use of DNA in the criminal justice system.

The legal system has played a significant role in influencing the implementation of QA measures in forensic DNA typing laboratories. In 1989, in *New York v. Castro*, Judge Sheindlin ruled that inadequate quality assurance was sufficient to hold DNA evidence inadmissible (see Case Citations - Quality Assurance). In November of 1989, the Minnesota Supreme Court, deciding *State v. Schwartz* (see Case Citations - Quality Assurance) became the first appellate level court to reject the use of DNA evidence analyzed by a forensic laboratory. The court was concerned that the testing laboratory had admitted having falsely identified two out of 44 samples from a proficiency test study performed by the California Association of Crime

Laboratory Directors (CACLD) (Kuo, 1988), and that the laboratory had not used relevant protocols developed by the FBI. These rulings, which were directed at private, commercial laboratories, reinforced the importance of quality assurance to the laboratories involved, and to public sector laboratories which generally moved more cautiously in bringing DNA testing results into court. Since these rulings, however, forensic DNA laboratories have recognized the importance of QA and few cases, if any, have resulted in DNA evidence being held inadmissible due to quality assurance deficiencies (Kirby, 1990).

The Federal Judicial Center's 1994 Reference Manual on Scientific Evidence (McKenna et al, 1994) written to help federal judges interpret and manage scientific evidence, includes a chapter on DNA evidence. Among various recommendations, it suggests the court inquire if the pertinent testing laboratory demonstrated an appropriate record of proficiency and quality control, "to permit confidence that the tests were conducted properly" (McKenna et al., 1994). If there is an adequate history of repeated proficiency tests then admission of laboratory error rates are appropriate [citing the U.S. Supreme Court's decision in Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993)]. The authors suggest courts may want to give more weight to blind test than open test results, and to consider the most recent error rate estimates, since the laboratory may have corrected earlier problems.

C. History of Clinical Laboratory Regulation

1. Introduction

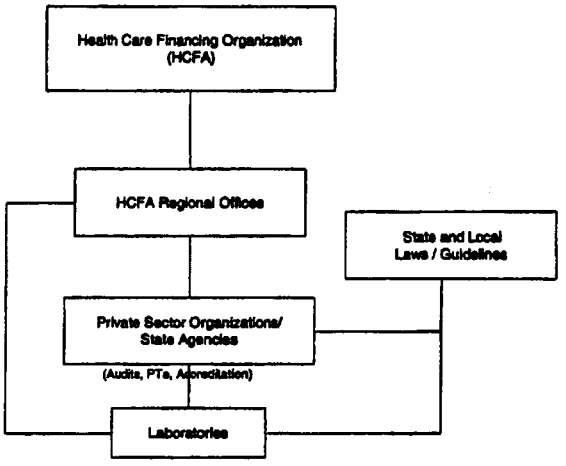
Currently, there are two federal regulatory programs for clinical laboratories, both administered by the Health Care Financing Administration (HCFA) of the Department of Health and Human Services (DHHS). The two programs are Medicare and Medicaid certification of facilities receiving reimbursement under these programs, and licensure of all clinical laboratories under the Clinical Laboratory Improvement Act of 1988 (CLIA '88). Please refer to Table II-1 at the end of this section for comparison of federal regulatory programs. In addition, the Food and Drug Administration (FDA) is another agency responsible for licensure and registration of facilities preparing, collecting, and shipping blood and blood products and for the approval of medical devices. The FDA activities are primarily directed toward a manufacturer's product, rather than the procedures and results of testing and the qualifications and activities of analysts (Edinger, 1988).

2. Medicare/Medicaid Regulatory Program

Federal authority over clinical laboratories began with the passage of the Social Security Act of 1965 (P.L., 89-97). This law established a system for the payment of benefits for medical care for several categories of individuals, including the aged, financially needy, dependent children, and the disabled. HCFA has primary responsibility for the administration of the Medicare program and for the provision of assistance to the states for the administration of the Medicaid program. Under this law, facilities must be approved to be eligible to receive reimbursement under the Medicare program. Under the Medicaid program, the states can impose additional requirements.

Figure II-1

Medicare/Medicaid 1965-1969



The Medicare regulatory programs are based on standards developed by the Secretary of DHHS to assure the health and well-being of individuals being provided with health care in a variety of inpatient and ambulatory settings including clinical laboratory testing. These regulations consist of several components including standards for personnel, record keeping, management, fire safety, internal quality control, and external quality control (proficiency testing). These facilities must also be in compliance with state and local laws including personnel licensure, facility licensure, fire safety requirements, and other related health and safety laws.

3. Medicare Regulatory Process

The Medicare regulatory structure is based on the models of several states (e.g., New York and California) and private sector programs (such as the College of American Pathologists [CAP] and the Joint Commission on the Accreditation of Health Care Organizations [JCAH]) and, therefore, shares many elements of these programs (See Figure II-1).

The Medicare certification and decertification process begins with recommendations from Medicare state agencies under contract with the HCFA to the HCFA regional office. Facilities in compliance with the standards are recommended to the HCFA regional office. The HCFA regional office makes the final decision whether or not to approve the facility to participate in the Medicare program. Facilities must also meet the same standards to be approved to participate in the Medicaid program. If the facilities do not meet specific standards, payment is denied. An independent laboratory may be denied payment in any specialty or subspecialty, or the entire laboratory may be decertified and not allowed to participate in the program.

4. Clinical Laboratory Improvement Act of 1967

The Clinical Laboratories Improvement Act of 1967 (CLIA '67), or the interstate licensure program, originated with the passage of the Partnerships for Health Amendments (Public Law 90-174) and was based on the decision by Congress to assure the quality of the testing performed on specimens in the course of interstate commerce. The program, unlike the Medicare program, was not limited to recipients of federal benefits programs, but applied to all laboratories whose testing could be considered interstate commerce. CLIA '67 adopted the Medicare personnel standards and added quality control and proficiency testing standards (Department of Health, Education and Welfare [DHEW], 1967). In addition, the interlaboratory surveys designed for educational purposes became the basis of licensure programs required for laboratories under CLIA.

It has been the contention of some clinicians that CLIA '67 was enacted primarily in response to misrepresentations of poor laboratory performance to legislators and to the public. In 1967, at hearings before the Committee on Interstate and Foreign Commerce of the House of Representatives, the Secretary of Health, Education, and Welfare testified that a 25% error rate was common among clinical laboratories (Boeckx, 1992). This statement was refuted by many scientists who discovered that the Secretary's report included information that was 20 years old. In addition to this congressional testimony, the New York Times, as well as many other media sources, published stories which featured the errors committed by clinical laboratories (Boeckx, 1992).

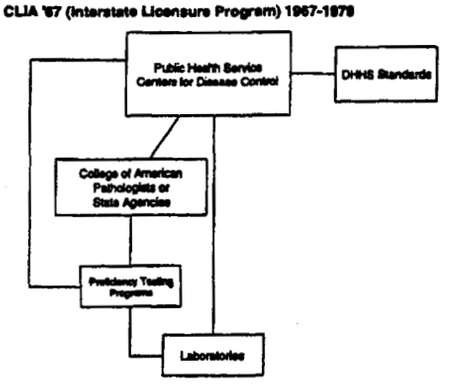
It appears that the federal government had decided that proficiency testing would become the "gold standard for judgment" in laboratory competency (Glenn, 1988). The new laws required proficiency testing for all governmental agencies and private sector organizations concerned with laboratory regulation and accreditation (Peddecord and Hammond, 1990). The laws made federal reimbursement for laboratory testing contingent on acceptable performance in federally approved PT programs. The impact of this legislation was revolutionary in that it altered the view that these proficiency testing programs were designed basically for self improvement and self-education (Glenn, 1988).

The CLIA '67 statute provided for the administration of the interstate laboratories program by the Public Health Service (PHS) (See Figure II-2). The Centers for Disease Control (CDC) was given the responsibility for the implementation and administration of the programs within the PHS. The CDC performed direct federal surveys of all interstate laboratories in which PT was to serve as the focal point for assurance of the quality of services. Interstate laboratories were required to enroll and successfully participate in the CDC's proficiency testing program for each test performed in interstate commerce and for which the CDC offered a program. The statute also required additional standards to be developed for internal quality control, personnel, and record keeping (DHEW, 1967).

The Medicare statute, in contrast, had no specific directive as to what standards were to be developed for laboratories. Under CLIA, the notable difference in the regulatory process is successful participation in the CDC proficiency testing program. There was a specific grading scheme in the regulations for the CDC PT program. In the case of Medicare, there were no federal grading criteria in the regulations, nor a definition of what constituted successful performance in proficiency testing for individual analytes or organisms. In the mid-1980's, the CDC curtailed proficiency testing, in lieu of its own, approved several state and private sector programs.

The CLIA program regulations provided for a grading system of individual analytes and provisions to revoke licensure if laboratories failed against an individual analyte test.

Figure II-2



Unlike the Medicare program, CLIA licensure actions only take place after a facility has had the opportunity for a hearing. The Medicare program was set up to take actions on specialty and subspecialty areas. Specialties and subspecialties are based on traditional laboratory practice and tend to involve similar technologies. Because Medicare and CLIA didn't always categorize tests in the same specialty/subspecialty, there was the potential for a laboratory to have its CLIA license revoked for a test (due to proficiency test failures) and still be reimbursed for the same test under the Medicare program (DHEW, 1967).

5. Consolidation of the Federal Regulatory Programs

Between 1967 and 1979, there were separate instances of overlapping federal clinical laboratory regulatory programs (Macro Systems Inc., 1986). CDC performed direct federal surveys of interstate labs (except those exempt because of NY and Wisconsin state licensure, CAP accreditation, or low volume); Medicare performed surveys of independent and hospital laboratories using Medicare State Survey agencies as well as JCAH and AOA accreditation programs; and the FDA conducted direct federal inspections of blood banks and transfusion services which, in most instances, also were part of Medicare and CLIA-inspected facilities. A number of CLIA labs were also in Medicare programs. By 1974, the Medicare program had modified its requirements and adopted CLIA '67's QC and PT standards.

In 1979, the PHS and HCFA signed an interagency agreement that was approved by the Secretary of the DHHS. The agreement consolidated the administration of the Medicare and CLIA laboratory programs within HCFA. The HCFA was responsible for the survey and certification and/or licensure of all clinical laboratories in both programs. The HCFA also assumed responsibility for taking adverse actions against laboratories, including denying licensure or certification, initiating action to revoke or suspend CLIA licensure based on failures in the CDC's PT program, and recommending any necessary criminal actions based on violations of the CLIA statutes.

6. Unification of CLIA and Medicare Programs

More recently, the HCFA was given responsibilities for developing new regulations and the CDC was given the responsibility for technical input and developing advances in proficiency testing. Together these agencies worked to develop a mandatory proficiency testing program which would be included in the revised regulations of the Clinical Laboratories Improvement Act of 1988 (CLIA '88) (P.L., 100-578) and the new HCFA requirements of 1990 (HCFA, 1990a). The modifications include: reliance on PT as a measure of quality; developing and modifying new quality assurance standards for newer specialties; elimination of dual and sometimes conflicting categories for licensure and certification; and revision of personnel, PT, QA management, program administration and hearing procedures.

In a fashion similar to what led to the passage of CLIA '67, it has been the contention of clinicians that CLIA '88 was passed due largely to public and legislator misperception of laboratory performance. In particular, media stories highlighting laboratory quality problems in the areas of cervical cytology screening, AIDS testing, and drug screening served to arouse Congressional concern (U.S. Congress, Hearings, 100-32; 100-529; 100-43; 100-765; 100-146). It was also thought by some that CLIA '67 was too lenient and that penalties needed to be increased (Laessig and Ehrmeyer, 1992).

Although the media sensationalized the issue, there were many legitimate concerns. The foremost concern was that there was no uniformity between Medicare and CLIA proficiency testing programs. It was necessary to provide some alignment of regulatory requirements among laboratories engaged in interstate testing, Medicare labs, unregulated labs, physician's office laboratories, and state regulated labs. Given this patchwork of regulation, the federal government was the logical entity to align all the disparate requirements. Similarly, it became clear that the mechanism needed to apply uniformly to all laboratories.

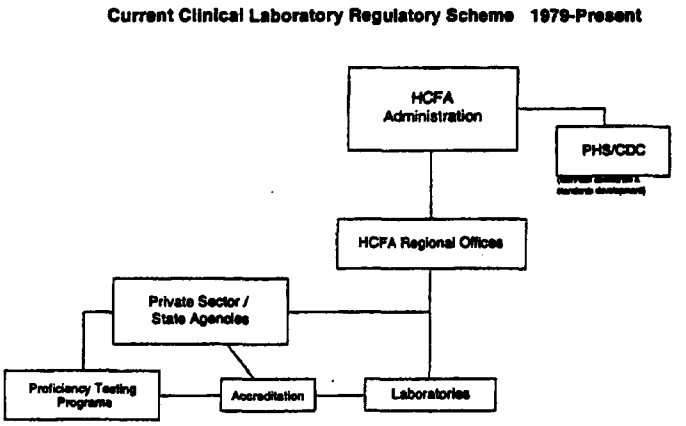
During this time, Medicare laboratories were not subject to the same national-proficiency-testing program as laboratories under CLIA '67 but were required to participate in state- or Secretary-approved PT program for each of the specialties or subspecialties of services offered. In addition, the HCFA had not established minimally accepted requirements in terms of program content, challenges, frequency of test events, and grading criteria. Each state was required to develop criteria for an acceptable PT program for the federally regulated laboratories in its jurisdiction. The federally issued State Operations Manual (HCFA, 1986) provided a list of Secretary approved programs but only minimal guidance to state agencies for implementing proficiency testing in these laboratories. Therefore, at the time, there were no consistent criteria from state to state. Satisfactory performance criteria could have ranged from 70% to 100% from subspecialty to subspecialty, but also varied from one state to another. As a result of these inconsistencies in monitoring proficiency testing, pass/fail standards, and grading criteria, certain affiliated laboratories operating in different states would sometimes find themselves involved in an adverse action in one state, as a result of PT results, but under no adverse proceedings in another state for the same testing scores. Some states also required enrollment in PT but did not monitor results (Hinkel, 1992).

7. Clinical Laboratory Improvement Act of 1988

In response to the problems of the existing regulatory scheme, the HCFA, CDC, state health officials, various private-sector organizations, concerned members of the laboratory industry, and the public, decided to consolidate all of the CLIA '67 and Medicare/Medicaid laboratory requirements (See Figure II-3). HCFA published a proposed rule in August 1988 to

revise CLIA '67 (HCFA, 1990b) and in October 1988, Congress enacted CLIA '88 (P.L. 100-578). This new regulatory model was based on four principles: personnel standards, quality control, quality assurance, and proficiency testing.

Figure II-3



A new set of guidelines, consistent with the standards established under CLIA '88, was passed on March 14, 1990. The rules were called "March 14, 1990 final rules" under CLIA '67 (HCFA, 1990a). These new rules defined explicit grading practices and what constituted acceptable laboratory performance. The "final rule" has formed a framework for establishing the CLIA '88 standards for quality control, quality assurance, record keeping, and proficiency testing. The proposal was published as a "final rule" and constituted current regulations for laboratories performing interstate testing and/or participating in the Medicare/Medicaid programs. In fact, CLIA '88 regulates virtually all clinical laboratories. Under CLIA '88, laboratories are regulated by either specialty, subspecialty, analyte, or individual tests. The privilege to perform tests is heavily weighted toward successful participation in proficiency testing of those areas for which the laboratory is seeking licensure or certification.

8. Proficiency Testing under CLIA '88

Most notably, CLIA '88 mandates proficiency testing of all clinical laboratories (HCFA, 1990a). The House conference committee wrote that it believed in "...determining a laboratory's competence (through proficiency testing) since it purports to measure actual test items" and "that proficiency testing should be the central element in determining a laboratory's competence (HCFA, 1990a)." In other words, the regulatory process to evaluate the quality of the laboratory performance is primarily based on proficiency testing results.

CLIA certified and Medicare-approved labs are required to enroll in DHHS approved proficiency testing programs for each specialty and subspecialty of service for which they seek certification (HCFA, 1991). Currently, there are approximately nineteen approved PT programs. When a laboratory wishes to be certified in an area, it must notify HCFA of the appropriate PT program chosen, participate in it for one year in a routine manner, and sign a form (with exception in the area of cytology) attesting that they did not give the PT sample any different treatment than routine patient samples.

In general, proficiency testing programs will provide five samples for each analyte or test three times per year (i.e., three testing events per year). After the laboratory has tested the

samples, the results are graded by the PT provider. The results are compared with the consensus answers from referee laboratories for the same specimens. With few exceptions, the passing score is 80%. If a laboratory receives a failing score on a PT, the laboratory must take necessary actions to find, correct, and document any problems in the testing performance. If the lab fails to perform acceptably for a given subspecialty (e.g., failure in two out of three PT events), CLIA certification will be terminated and Medicare disapproved, but only for that subspecialty. In addition, if it is determined that a laboratory has intentionally referred a proficiency test specimen to another laboratory, the laboratory will lose its Medicare approval and will have its certification revoked.

Recognizing the shortcomings of this assessment method, CLIA '88 "...is designed to strengthen and improve proficiency testing." For example, regulatory or accreditation systems that use PT as a performance criterion recognize that single failures are not uncommon. Under CLIA regulations (which are also adopted by accreditation agencies certified as meeting CLIA requirements by the HCFA), a laboratory must fail either two consecutive PT events or two of three events before a regulatory response occurs. The new rules for proficiency testing are more explicit than the guidelines found in CLIA '67. For example: a quarterly PT with uniform grading system was established; authority was given to the government to conduct on-site PT to assure that samples were handled without special treatment; and PT results were to be made available to the public. What was also explicitly written into the law was specific language concerning proficiency testing and PT programs. This included: enrollment in a program; testing of samples; successful participation; procedures for reinstatement after proficiency test failure; and approval and disapproval procedures for PT programs. Enrollment and successful participation in proficiency tests by laboratories are conditions in the new rule. The current requirements emphasize the increased importance of evaluating and achieving a passing score on specimens of known content, which are to be processed/tested as if they were patients' samples and serve as a measure of laboratory quality.

In addition, each specialty and subspecialty have requirements unique to that area (e.g., cytology is unique within CLIA '88 in that it is the only laboratory specialty for which Congress wrote specialty specific standards [1, Sec. 353(f)(4)(B)]; in compatibility testing, failure of attaining a score of at least 100% in a proficiency testing event is unsatisfactory while in most other specialties, scores less than 80% is considered a failure) (HCFA, 1990a; Laessig and Ehrmeyer, 1992; Lanphear et al., 1992). With the passage of CLIA '88, standards were also developed in newer specialties such as cytogenetics, DNA probes, molecular genetics, and updating standards for other areas ranging from histocompatibility testing to cytology. Currently, CLIA '88 requirements apply to all 150,000 hospital, reference, physician, and clinical laboratories in the United States.

Table II-1 Comparison of Clinical Regulatory Programs

Medicare/Medicaid Program	CLIA 1967	CLIA 1988
o designed for laboratories receiving Federal reimbursement for laboratory tests o no consistent grading of PT (varies from state to state) o licensure by specialty / subspecialty	o designed for interstate commerce laboratories o PT grading criteria is standardized nationally o PT surveys are the basis of licensure	o regulates all clinical laboratories with few exceptions o more explicit guidelines regarding PT o licensure by individual tests, analyte, specialty and subspecialty o criteria developed for PT providers o standards developed for newer specialties

D. Quality Assurance / Quality Control and Proficiency Testing in Clinical Laboratories

The intent of a QA/QC program is to improve laboratory performance through a combination of: (a) self assessment; (b) self improvement and education; (c) quality assurance/quality control program; (d) accreditation; and (d) regulation. A laboratory may assess its current performance through a variety of methods including internal QC, internal audits, and most notably proficiency testing. After determining the areas of deficiency, laboratory directors can incorporate corrective and preventive actions such as QC methods, training, and/or education programs toward those areas to improve laboratory performance. Another important aspect of a laboratory's QA/QC program is that it often adheres to standards and guidelines created by specialty organizations or committees (i.e., Virology Committee of the AIDS Clinical Trial Group; National Institute of Allergy and Infectious Diseases-Quality Assurance Program; American Association of Blood Banks; American Society of Histocompatibility and Immunogenetics; British Society for Haematology; British Blood Transfusion Society). Most of these organizations have developed proficiency testing procedures in their standards (Growe, 1996)(Boulton et al., 1987)(Yen-Lieberman, 1996). Some laboratories have taken the initiative to develop their own internal blind proficiency testing programs. One such program has been established by a clinical microbiology laboratory in which testing is administered with progressive levels of organism identification difficulty (Estevez, 1980). Most notably, accreditation by a private sector organization (such as the College of American Pathologist's Laboratory Accreditation Program, or the Joint Commission on Health Care Organizations) is an indication of a laboratory's commitment toward quality and laboratory improvement. Therefore, successful PT results and subsequent accreditation have symbolized that a laboratory can offer quality results and ensure quality patient care (Hodnett, 1999). The federal government has published regulations and requirements that a clinical laboratory is required to uphold. Under

current regulatory processes, a laboratory must be licensed or certified to perform certain tests. It is worth noting that the federal government recognizes laboratories that are accredited by organizations with so-called deemed status. A laboratory accredited by an organization with deemed status is equivalent to a laboratory that has been licensed or certified by designated government agencies.

The first PT programs were voluntary and based on the analysis of prepared solutions whose contents and/or concentrations were unknown to the analysts. The first proficiency testing program began in 1946 when Sunderman distributed anonymous specimens to hospital laboratories to assess laboratory performance and to standardize results. They found that results varied widely among laboratories (Belk and Sunderman, 1947). Through interlaboratory comparison and consensus results, laboratory directors could assess the quality of work performed in their laboratories and ascertain the causes of analytical discrepancies.

During the 1940s, the College of American Pathologists (CAP) was founded and instituted the first national proficiency survey called the Standard Solutions and Materials Program (Wagner, 1997). This program was similar to the one conducted by Belk and Sunderman. In general, PT programs were viewed primarily as a mechanism for continuous, incremental improvement process. It was this underlying philosophy that led CAP to organize, promote and, in essence, to mandate proficiency testing as a criterion for laboratory accreditation. CAP accreditation was not based exclusively on acceptable proficiency testing results, but also on the successful evaluation of personnel standards, QC checks, internal/external audits, education, and training. Gradually, it became recognized that in order to maintain high standards, the accuracy of laboratory measurements must be under constant professional surveillance (Sunderman, 1992). Although participation in proficiency testing was voluntary, by the late 1940s and early 1950s, compulsory PT was required by some professional societies, as well as some state and municipal governments. In the 1960s, proficiency testing became a standard practice in clinical laboratories. By 1961, the CAP had become what is now the largest voluntary peer review program in laboratory medicine. The CAP hoped that this voluntary accreditation program would help moderate the impact of anticipated regulatory control (Rej and Jenny, 1992).

Simultaneously, internal QC efforts began to play an increasingly central role as laboratories perceived a need to ensure their analytic performance. Samples derived from stable serum pools were repeatedly analyzed to generate numerical data that set limits for acceptable daily performance. The serum pools were shared among many laboratories and the results were statistically reviewed by various organizations to provide both group and individual laboratory analytic limits of acceptable performance (Glenn, 1988). The first program of this type was developed in 1967 by Joseph A. Preston and involved 30 cooperating laboratories in the Colorado area (Copeland and Rosenbaum, 1972).

In the 1960s, interlaboratory surveys had become more efficient and comprehensive. The clinical laboratory community developed confidence in PT based on the following improvements (Laessig and Ehrmeyer, 1988):

(a) "large pools of uniform quality assurance specimens could be produced and preserved;

(b) data could be processed by large computers to yield compilations comparing individual results with the group's consensus, mean value, or true value if known;

(c) surveys had been able to demonstrate if methods in routine laboratory use were accurate or possessed a known, measurable bias;
(d) laboratory directors, when reacting to interlaboratory surveys, had demonstrated the ability to correct errors quickly;
(e) schemes, based on state of the art measurements of interlaboratory imprecision, were available to grade results and to assist laboratory directors in initiating corrective actions."

These advances in the laboratory have been reflected in the dramatic improvement of PT results from the 1940s to the late 1960s. In 1969, a retrospective look at 6 years of performance on CAP surveys was conducted by Skendzel et al (1970). They observed that the coefficients of variation had narrowed by 50% or more for all analytes Belk and Sunderman had studied, except for cholesterol. Other evaluations conducted more recently have confirmed that the large variances among individual laboratory's results have diminished.

With the passage of both CLIA '67 and '88, the regulatory process mandates proficiency testing of all clinical laboratories. The fundamental premise of mandatory proficiency testing is that by requiring all clinical laboratories to participate in an acceptable program, quality results will be ensured (Laessig et al., 1992b). An opposing view is that if the proficiency testing process focuses principally on laboratory performance, the primary incentive to pass the proficiency tests is to gain licensure (Laessig and Ehrmeyer, 1988). Nevertheless, three decades of federal clinical laboratory standards have been associated with improved laboratory performance. In fact, it has been demonstrated that mandated proficiency testing enhances overall quality of clinical laboratory testing, including turnaround time, accuracy of results, and training of laboratorians while voluntary or self regulation have been found to be less effective in achieving goals (Boone, 1993) (Crawley at al., 1986). It has been acknowledged that proficiency testing does have limitations and that it may not be the optimum tool for laboratory regulation; however, many believe it is the single most cost effective means available to regulators for evaluating laboratories (Laessig et al., 1992a)(Bartola, 1988). Federal policy itself may not be capable of assuring quality performance in laboratories but it can be a powerful engine mobilizing resources to promote quality (Hammond, 1988).

Under CLIA, proficiency testing is a regulatory requirement; however, the clinical community still views PT as just one component of QA that is intended to improve laboratory and analyst performance along with education, training, and methods development (Boone et al., 1985) (Salkin, 1997). It is worth noting that proficiency testing has also become an important tool for the evaluation of new technologies and/or new protocols such as flow cytometry cross matching and DNA probing (Scornik et al., 1997)(Dewald et. al., 1997)(Dewald et al., 1996). Currently, proficiency testing extends itself into almost all activities of the clinical laboratory including clinical chemistry, hematology, microbiology, immunology, anatomic pathology, and newer specialties including cytogenetics, and molecular genetics. Although CLIA's critics suggested that the mandated quality standards for PT would lead to a higher incidence of failed laboratories, it has not proven to be true; on the contrary, the inspection and PT performance data has strongly shown that overall quality of laboratories has been improved, which indicates that CLIA's mandated quality performance and standards for PT are achievable (Ehrmeyer and Laessig, 1999).

E. Limitations of Proficiency Testing

Compared to other forms of quality control checks such as internal and external audits, open proficiency testing is a relatively inexpensive method for assessing laboratory performance (Boone, 1988). Shahangian (1998) performed an extensive review of clinical laboratory proficiency testing and determined that there is a positive relation between PT performance and other putative quality indicators of laboratory performance; however, there are limitations to the usefulness of PT data. Proficiency testing data are an indicator, but not a definitive measure of laboratory performance. The limitations of PT practices are: (1) incomplete testing of the total testing process (TTP); (2) special treatment of PT materials; (3) the "matrix effect;" and (3) how PT performance criteria are utilized.

1. Incomplete Testing of the Total Testing Process

Because proficiency testing materials originate from a different source than patient specimens, PT samples enter the testing process at the late pre-analytical phase of the total testing process (TTP) rather than at the beginning of the TTP. Therefore, open proficiency testing only assesses the analytical stage of the TTP.

2. Special Treatment of Proficiency Test Materials

It is also generally believed that laboratories give special attention to known or open PT specimens and, therefore, the results from a proficiency test are not truly representative of routine performance of a laboratory. In a survey conducted by Cembrowski and Vanderlinde (1988), it was determined that various practices were used by laboratories to improve performance on PT specimens. These practices included replicate analysis, sending the PT sample to a designated analyst, analyzing PT specimens immediately after standardization and quality control, and delaying analysis until the analytic process was optimal. It is on the basis of findings like these that proficiency test results are often thought to be the best a laboratory can produce. (It should be noted that these issues have been addressed under CLIA '88 guidelines--laboratories must attest to the fact that no special treatment was given to proficiency test samples).

3. Matrix Effect

Typically, proficiency test specimens are manufactured samples that simulate patient specimens and because they are not the same, PT results can be difficult to evaluate and control (Duckworth, 1988). Specifically, contributions to error are the confounding effects of "fluid-matrix" caused bias, method instrument bias, and deviations from methods associated with analyzing PT specimens. Proficiency test specimens are suspended in solutions (a "fluid-matrix") to approximate clinical and biological conditions. Clinicians have observed a "matrix effect" in which the fluid-matrix may destabilize the PT specimen over time and/or cause interference in instrument readings. Thus, the true value of the proficiency test may be biased. With today's ever changing technologies, timely and continuous expert input is required to identify, correct, and/or account for such things as matrix interference and method instrument bias.

4. Proficiency Test Performance Criteria

A PT program should distinguish between good and deficient laboratories. As such, proficiency testing evaluation criteria must be used to determine whether or not a laboratory is performing acceptably. Aside from focusing only on the analytical stage of the TTP, some of the limitations of setting performance criteria are the fewer numbers of PT samples tested compared with patient samples, and difficulties in establishing "true values" for samples. It is important to establish true values to ensure the basis of uniform standards in clinical laboratories. In response to these potential problems, the clinical laboratory community has developed four basic means by which proficiency test target values may be established: (a) consensus values or peer group statistics after appropriate outlier exclusion, submitted by participating laboratories; (b) analysis of specimens by definitive methods or protocols correlated to definitive methods; (c) referee laboratories; and (d) documentation of the composition of the specimen by design and method of manufacture, by the manufacturer.

The most common PT grading criteria based on consensus values is the use of peer group means and specified limits (usually 2 or 3 SD from the peer group mean) or a specified percentage or interval from the target value or "true value." One of the disadvantages of using peer group statistics is that they do not take into account systematic or random errors specific to the methodology (Rej and Jenny, 1992). In addition, it is possible that laboratories might standardize and calibrate on a biased consensus result. In order to produce a more accurate assessment of the overall quality of laboratory testing within an interlaboratory survey, a proficiency testing program should involve participation by a wide spectrum of laboratories representing all levels of performance.

As an example of (b) above, the National Bureau of Standards (NBS), [now known as the National Institute of Standards and Technology (NIST)] has recognized the value of assigning definitive values (as opposed to consensus values) to the analytes in interlaboratory surveys and has determined the practical analytic goals for accuracy. External quality control programs that use accuracy based target values and fixed evaluation limits can provide a technological basis for improving interlaboratory accuracy. These definitive values, or target values, may be determined through the use of exacting protocols, state of the art equipment, and methodologies (Welch and Hertz, 1988). Developing new methodologies, however, is not without its disadvantages. These methods are often slow, tedious, costly and may necessitate development of new techniques (Glenn, personal communication, 1996).

The methods for the analysis of a particular analyte may differ from laboratory to laboratory. Because of the variety of methods and/or instrumentation available to laboratories, clinicians have been concerned with the uniformity of proficiency test results and the possible effects on a survey. In a study conducted by AuBuchon (1991), all the analytical methods for the analysis of alanine aminotransferase (ALT) were compared. With regards to test results, only small differences were found among analytical methods. Nevertheless, the NIST has developed definitive methods for use in testing certain analytes. The number of analytes for which definitive methods are available is small; however, for those analytes that do not yet have definitive analytic methods, survey-verified grand consensus mean values often come very close to true values as shown by Ehrmeyer and Laessig (1988), Gilbert (1978), Grannis (1976) and Hartmann et al (1985).

F. Factors Relating to Proficiency Test Performance

Performance in proficiency testing programs has been shown to be related to several laboratory characteristics including duration of participation in a PT program, personnel qualifications, laboratory environment, testing methodologies and automation, and QC procedures.

1. Duration of Participation

A long accepted quality assurance maxim holds that "anything improves if you measure it" (Hammond, 1988). The notion that performance in proficiency tests improves over time is generally accepted. In fact several studies have been conducted to determine whether rates of unacceptable results decrease as laboratories gain experience in interlaboratory comparison programs. Hansell and Haven (1979) first showed long-term improvements in interlaboratory agreement in the CAP Ligand Assay surveys from 1972-1978. In Data ReCAP, 1970-1980, researchers for the CAP showed that interlaboratory agreement improved markedly for most analytes over time; improvement was particular striking in the first two years of a PT program (Elevitch, 1981). More recently, data was examined from CAP surveys from 1987 to 1993 (Tholen et al., 1995) in the areas of chemistry, hematology, immunology, and blood banking. The authors found that laboratories with consistent participation show consistent and statistically significant improvement in performance for the first 3 to 4 years of proficiency testing and that laboratories with more experience in proficiency testing have lower rates of unacceptable results.

2. Personnel Qualifications

More experienced, better trained analysts, greater specialization of laboratory workers as well as certification of technologists are also related to PT performance. CLIA '88 called for an assessment of personnel competency. In response, a review of personnel standards was conducted in 1996 by Peddecord et al. in which the relationship between laboratory personnel regulations and laboratory performance was examined. By utilizing proficiency tests results as a measure of lab performance in relation to personnel regulations, it was determined that higher PT results were usually associated with higher personnel qualifications. The study also examined the concepts of cost-effective analysis (dollars expended per health outcome attained), cost-benefit analysis (dollars expended per dollars of benefit achieved by placing value on various states of being) and cost-utility analysis (dollars expended per life years saved) of personnel regulations. The paper has described these as being difficult to perform due to the fact that there is no standard method to do this. Finally, the issue of competency assessment was addressed. Some of the factors include technical competence, professionalism, and morals and ethics. The researchers concluded that such an analysis would also be difficult to perform due to differing opinions of competency. Jenny et al.'s (2000) article on the causes of unsatisfactory performance in proficiency testing of toxicology laboratories provided several insights in identifying causes of laboratory error. Guidelines were produced to assess laboratories in deciphering the causes of spurious results and what was termed "common-cause analytical error." Spurious results were often caused by misinterpretation of instrument codes and mishandling of data on instrument printouts. Common-cause analytical error was most commonly attributed to systemic error and to "calibration drift (error)." The PT provider plays a critical role in identifying common causes of error among participants and sharing this information with manufacturers of instrumentation.

Laboratory tolerance of error must be informed by the performance limits of the instrumentation itself.

3. Laboratory Environment

It has been demonstrated that certain types of laboratories, specifically hospital and independent laboratories, tend to perform better in proficiency tests than physician office laboratories. Belsey and Baer (1986) observed a PT passing rate of 95% in hospital and independent laboratories as compared with a rate of 69% to 86% for physician office laboratories. Stull et al. (1998) observed that the aggregate rates of satisfactory event performance for all regulated analytes, tests, and specialties were 97% in hospital and independent laboratories and 91% in all other testing sites. The hypothesis for worse performance in non-hospital and independent laboratories is that these laboratories were previously unregulated and not required to take proficiency tests until 1994, and, therefore, probably reflects lack of laboratory expertise and/or experience with proficiency tests.

Better PT performance has also been positively associated with increased test volume for certain analytes such as cholesterol (Erickson et al., 1991) and glucose (Phillipou et al., 1990). Shehangian (1998) determined that increased institutional size and laboratory workload have also been generally related to improved PT performance and less variation in chemistry, bacteriology, parasitology, and qualitative hematology.

4. Testing Methodology and Automation

Another explanation that should not be overlooked is advances in testing methodology and automation. In physician office laboratories, automation was related to increased precision and reduction in error rate by a factor of 1.5 to 3 (Bloch et al., 1988).

5. Quality Control Procedures

Jones et al. (1991) described improved performance in the CAP Microbiology Surveys from 1972 through 1989 and attributed many of the improvements to interlaboratory standardization of methods. Positive relationships between better quality control practices and better proficiency test performance are also significant. Lawson et al. (1988) showed that PT results are related to measures of performance in a laboratory's quality control system. In addition, they determined that a laboratory's accreditation status can affect PT results. A review of microbiology laboratory performance over a 20-year period was conducted by the Ontario Medical Association's Laboratory Proficiency Testing Program (LPTP) (Richardson et al., 1996). This study, in general, has revealed that performance improved over time and has determined that poorly performing laboratories had common characteristics, including inadequate supervision, limited continuing education opportunities, lack of effective quality control, use of nonstandardized methods, poor documentation of work performed, and over investigation and over reporting of clinically irrelevant bacteria.

Two recent studies (Strand, 1994; Bloch et al., 1988) have concluded that improvement of laboratory performance was not the direct result of the PT process itself, but primarily because of two factors: (1) extensive education that was a key component of the larger QA/QC program; and (2) voluntary withdrawal from testing by laboratories displaying poor performance. Witte et al. (1997) studies more than 200,000 clinical chemistry results and sought to identify differences from expected values. Only 98 results were deemed unacceptable and, of these, only 9 judged to potentially cause error in patient management. The authors compared such unacceptable rates

with other industries like electronics, physicians office laboratories, and HIV testing in the military. Whereas earlier authors attributed the majority of errors to sample mix-ups, the present authors believe the unacceptable performance was possibly the result of malfunction of automated analytical equipment.

G. Other Examples of Quality Assurance / Quality Control and Proficiency Testing

QA/QC and proficiency testing have also played important roles in non-clinical fields such as environmental and industrial testing. These laboratories test in a wide range of areas including water bacteriology, wet chemistry, trace metals, lead testing, pesticides and food. In the same fashion as the clinical laboratory community, specialty organizations such as the Association of Official Analytical Chemists and the International Organization for Standardization have developed PT guidelines and QA/QC standards. Some of these PT programs include the American Industrial Hygiene Association' (AIHA) Environmental Lead Proficiency Analytical Testing Program (ELPAT) and the UK's Food Analysis Performance Assessment Scheme (FAPAS). One example of governmental regulation in environmental laboratories began with the passage of the Clean Water Act in 1972 and the Safe Drinking Water Act in 1974. This stimulated regulations administered by such agencies as the Environmental Protection Agency (EPA), the Food and Drug Administration (FDA), and the National Institute of Occupational Safety and Health (NIOSH). Federal regulation has also led laboratory accreditation by "deemed" accrediting organizations such as the New York State Department of Health, AIHA, and the American Association for Laboratory Accreditation (A2LA). Like the clinical laboratory, industrial chemical laboratory accreditation is awarded by appropriate organizations when QA/QC criteria as well as proficiency testing are fulfilled. In addition, certification is given to laboratories by approval category rather than individual test. In other words, a laboratory is only qualified to perform categories of similar tests based on results of proficiency testing. There are certainly many parallels, and researchers in this area have arrived at the conclusions similar to those the clinical laboratory community has reached--that proficiency testing, coupled with strong QA/QC programs and statutory authority provides the incentive for laboratories to maintain high levels of performance leading to improvement (Esche and Groff, 1997)(Key et al., 1997)(Thompson and Wood, 1993)(Daly and Asmus, 1985).

H. Comparing Blind and Open Proficiency Testing

Mailed proficiency testing samples have served as a basis for evaluating the performance of clinical laboratories for many years. These specimens are considered open, or declared, samples. In other words, the specimens are known to be PT specimens; however, the correct responses are unknown to the analyst. The proponents of open testing have recognized its value, as well as its limitations. Since most PT specimens are identified as such, the laboratory staff know they are being tested. Consequently, open PT and other routine forms of evaluation where the examiner knows he/she is being tested, are measuring ideal analytic capabilities rather than those under routine conditions. Prior to the passage of CLIA '88, many accepted that proficiency test specimens receive special treatment, and at least one study documented this observation (Cembrowski and Vanderlinde, 1988). Another study also documented evident collusion on proficiency tests among physician office laboratories in a small geographic area (Boone, Hearn and Lewis, 1985). As a result of such experiences, it has been suggested that if an unknown or blind PT specimen were submitted to a laboratory in the guise of a routine specimen

and is undetected, it could not receive special attention. Hence, the results of a blind PT would be a "truer" measure of laboratory performance.

1. Urine Drug Testing and Toxicological Analysis

Blind testing in forensic urine drug testing (FUDT) has an extensive history and literature. Currently, some proficiency testing programs, such as the Department of Defense's proficiency testing program for FUDT (Smith, 1990) and HIV testing, are blind. The first FUDT PT programs were conducted by the CDC and NIDA from 1972 to 1981. Nearly all of the studies comparing blind PT to open PT performance are based on survey data collected from the laboratories participating in the CDC and/or NIDA programs. The first studies were qualitative in nature and the assessment of laboratory performance was based on the percent of correct drug detections (Lamotte et al., 1977; Mason, 1981; Boone et al., 1982; Hansen, Caudill, and Boone, 1985). The blind PT samples (drug free urine specimens spiked with a variety of commonly abused drugs) were submitted as patient specimens to a group of FUDT laboratories.

The first comparative analysis of proficiency test results was conducted in 1976 (Lamotte et al, 1977) and looked at data from two blind vs open trials, occurring in 1973 and 1975. Participating laboratories were given a set of open PT specimens and, at the same time, were given an identical set of specimens submitted through hospital administrators or physicians and disguised as ordinary patient specimens. In every case for which comparative data were available, the laboratories detected a greater percentage of the drugs in the open samples than in the blind samples.

In a study conducted by Hansen, Caudill and Boone (1985), 13 laboratories were evaluated with blind proficiency test specimens from 1973 to 1981. Each year, 100 blind samples were distributed to each laboratory. These blind results were compared against CDC open PT test data from 1979 to 1981. The findings consistently showed that blind PT samples resulted in a lower correct response rate and higher rate of false negatives. In addition, in spite of advances in methodology, the results from blind PT did not improve significantly over time.

A study was completed that examines CDC proficiency test data from 1978-1980 in the areas of drug monitoring, drugs of abuse, chemistry profile, and blood lead (Boone et al, 1982). The study involved the comparison of three types of surveys the CDC had conducted: 1) on-site surveys in which trained personnel visited laboratories that had experienced performance problems in the quarterly mailed proficiency tests and reviewed the laboratories' analytical procedures by using carefully referenced samples to determine sources of errors and providing assistance in correcting them; 2) special assistance surveys and then telephone consultations were conducted to correct the problems; and 3) blind surveys in which carefully referenced samples were sent through normal patient sample acquisition routes to assess the actual day-to-day performance capability of the laboratories. Again, the authors found that blind PT scores were "27 percentage points lower than the mailed cumulative averages" and that "...each laboratory's blind proficiency testing performance was rated unacceptable (Boone et al., 1982)." In addition, the results of this study suggested that on-site surveys by trained lab surveyors and special mailed assistance surveys can be very effective in identifying the source of analytical errors.

Quantitative studies were also conducted using blind samples spiked with various concentrations of analytes (Mason, 1981; Davis, Hawks, and Blanke, 1988). The results were similar to the original qualitative studies in that a surprising number of laboratories failed in the identification of a compound or a false identification was reported. Similarly, a large number of

reported quantitative results were outside an admittedly arbitrary acceptable range of the target values ('2 Standard Deviations (SD) from the mean or Coefficient of Variation (CV)'15% from target value).

In 1981, a more comprehensive study was published which reviewed data for laboratories performing toxicological analysis (Mason, 1981). Both qualitative and quantitative blind proficiency test samples were examined. Along with FUDT PT, the results of proficiency testing surveys in the area of toxicological analyses of human sera were analyzed. In a particular forensic toxicology proficiency test survey, the investigator looked at data from the 1950s and the 1980s and discovered that the 1980s results "show variability not unlike that of twenty years earlier." Examination of CDC blind FUDT PT data from 1973 to 1980 reveals that correct responses ranged from 36-60% in comparison to 100% correct scores on open mailed PT and on-site surveys. In contrast, the study looked at open surveys from 1979 and 1980 and determined that "...participants scored almost as well as reference laboratories" and that "...identification was very high and slightly exceeded that of the referees." Examination of data from blind proficiency testing in quantitative toxicology reviewed results from analyte to analyte. The study determined that most of the acceptable responses (CV was no more than 15%) were for analyses of familiar analytes such as ethanol, phenobarbitol, digoxin, and theophylline. The study also looked at two laboratories that participated in two blind PT studies in 1980. The two laboratories averaged 64% and 41% incorrect responses. These errors were "almost entirely false negatives" (Mason, 1981).

In 1987, the American Association of Clinical Chemistry (AACC) evaluated the ability of subscribers to the AACC Toxicology Survey Plus program to assess accurately the presence or absence of five drugs of abuse which were cannabinoid, cocaine, morphine, methamphetamine, and phenylcyclidine (Frings, White and Battaglia, 1987). They determined that urine drug testing "can produce accurate results." The overall accuracy was 99.2%, the false negative rate was 0.8% and the false positive rate was 0.05%. In response to criticism that the PT samples were open, the AACC repeated the study in 1989 (Frings, White, and Battaglia, 1989) by supplying participating laboratories with blind specimens. The blind results were comparable to the open PT results and the overall accuracy was 97%, the false negative rate was 2.36%, and there were no false positives. Although there was a slightly higher rate of false negatives, the investigators once again concluded that urine drug testing "can be accurate."

Thus far, the results suggest that in general, laboratories perform much better when the staff members know they are being tested. The greater attention given the open PT samples versus the blind samples seems to account for the difference in performance. It has been postulated that for certain analytes, the employment of less sensitive testing could be the cause of lower correct response rates on blind tests (Hansen, Caudill, and Boone, 1985; Davis, Hawks, and Blanke, 1988). If a particular test is not specified, the laboratory may opt for a less expensive test or method which may be less sensitive. Almost all of the aforementioned studies support the desirability of blind proficiency testing.

2. Clinical Chemistry

Glenn and Hathaway (1979) examined data from a hospital chemistry laboratory's QC program in which specimens (originally derived from previously used quality control sera whose values had been established many times) were re-submitted as patient specimens. The study not only examined analytical results, but also pre- and post-analytical errors. These errors include: clerical errors; average turnaround time for statistical procedures; printer malfunctions; specimen

handling problems; distribution failures; failures to perform all requested tests; time stamping; lost reports; and specimens misidentified in laboratory processing. In contrast to the FUDT PT studies, the study found that the analytical results from the blind specimens were comparable to the open samples; however, the authors found that blind quality assurance was useful in detecting problems in the pre- and post analytical phases.

Reilly et al. (1996) tested 22 laboratories in the specialty of blood lead analysis by using 828 specimens, both blind and open, from cows and/or goats that had been fed lead salts. Of the 22 laboratories involved in the study, six (27%) had open proficiency test scores of more than 80% and blind proficiency test scores of less than 80% using CLIA '88 criteria for result acceptability. In addition, two (9%) of these six laboratories had open PT scores that were higher than blind PT scores by more than 35%. The results of the study implied that most laboratories' open and blind proficiency test performance were comparable; some, however, failed the blind PT but passed the open PT using CLIA grading criteria.

3. Human Immunodeficiency Virus (HIV) Testing

In the late 1980s and 1990s, several blind trials were conducted in the field of HIV testing (Damato et al., 1988; Peddecord et al., 1988a, 1988b, 1988c; Schalla et al., 1996). Comparable to the Glenn and Hathaway (1979) internal QC sera study, analytic errors were less of a problem than nonanalytic errors. In 1988, the Department of the Army evaluated HIV testing laboratories that were part of its total quality assurance program (Damato et al., 1988). Each of the participating laboratories was rated on eight criteria including open and blind testing. In addition, the laboratories participated in internal and external proficiency panels for a period of 12 months. The authors found that blind PT results were nearly as good as open PT results (99.6% correct response rate in blind versus 100% in open). Furthermore, out of 1098 blind samples, only four errors occurred (0.4%) and all were false negatives.

In 1992, a study was funded by the CDC and the Association of Schools of Public Health to develop a method for establishing a blind proficiency testing system for HIV testing (Peddecord et al., 1988a, 1988b, 1988c). The goal of this study was to determine the practicability of blind proficiency testing and assess the feasibility of expanding such a system. This study was the first blind study in which human sera was used. Analytic and nonanalytic issues were examined and the authors found that test results were of "...high accuracy and relatively few errors attributable to laboratory performance." The few analytic errors that did occur were all false negatives. The explanation for these false negatives was that less expensive (therefore less sensitive) tests were being used which resulted in decreased sensitivity. Nonanalytic issues examined included turn around time, charges for testing and report content analysis. The study concluded that blind PT had been most useful in identifying nonanalytic problems and that although blind PT does provide a more valid measurement of routine performance levels, the "complexity and expense limits blind proficiency testing as an external quality assurance tool."

Schalla et al. (1996) also conducted a blind versus open performance evaluation involving HIV detection. The method used in this study was the split specimen design, which involves splitting the specimen to the testing laboratory as a patient sample (blind PT) and the second split specimen was sent to the CDC. The CDC then split this specimen three ways. One of the three split samples was sent as an open PT to the same testing laboratory, one other was tested by the CDC (the reference laboratory) and the third was frozen and tested in the event that the both the open and blind sample tested by the target laboratory disagreed with the reference laboratory

result. Of the 6967 pairs of split specimens, there were 61 (0.88%) discrepancies between the reference laboratory and target laboratory. About half of these discrepancies were also attributed to different sensitivities of the test kits. Of the 25 inaccurate results (i.e., comparison of the results from the open specimens and its blind specimen counterpart) obtained by the testing laboratories, 14 involved blind samples only, 9 involved open samples, and 2 involved both.

4. Conclusion

In general, the results from many of the earlier comparative studies show that laboratories perform better in open proficiency tests than blind proficiency tests. In addition, blind PT seems to detect more problems, particularly false negatives. However, later studies (late 1980s to present) show that open and blind results can be comparable with little or no differences. Presumably, this is due to improved technologies and higher laboratory standards. It is also possible that the better performance observed in open testing may be analyte dependent and may not be statistically significant. Lack of statistical significance in some cases may be due to inadequate statistical power resulting from insufficient sample size. Recognizing the advantages of blind PT in assessing laboratory performance, Davis et al. (1988) concluded that blind proficiency tests "should be conducted whenever the logistics can be worked out by contractors for laboratory services, clinicians using laboratory services, and the laboratories themselves to assure the continuation of quality service."

The general consensus of the clinical laboratory community is that blind proficiency testing, on a national scale, would be difficult to administer (especially in laboratories with a low volume of case work). The providers would be faced with distributing samples to a large number of facilities (~150,000). In addition, the costs of a blind PT have been estimated to be up to ten times the cost of an open PT (Boone et al., 1982). Even more problematic are the quantity and different types of samples that must be manufactured; i.e., given the number of challenges required by CLIA, a sufficiently large sample would be needed to establish a target value and to ensure all participants received standard samples exposed to the same environmental variables so that submitted results could be evaluated fairly.

Another potential problem is the possibility the blind proficiency test specimen will be recognized by the analyst. Changes in routine data handling protocol due to shortened data handling deadlines could potentially reveal the identity of the sample to the analyst. In addition, manufactured specimens used in QC activities may be recognized by their appearance, thus, may be detected as a PT specimen by analysts. A logistical problem with a distribution of blind proficiency test samples is the possibility that patient samples selected for blind specimens may be infectious. This could create an additional hazard for QA personnel and analyst alike. Untested sera or known infectious sera cannot be sent by common carriers unless identified and labeled as such.

Evident in the studies and trials conducted, a blind proficiency testing program can be successful but most likely on a small scale. Clinicians agree that a good compromise between an external open and external blind PT is intralaboratory blind testing. Essentially, this is a voluntary program in which a laboratory would conduct its own internal blind PT. There are many ways to conduct this type of program, including: (a) random re-analysis of a patient sample; (b) introduction of a test sample as a routine specimen; or (c) introduction of a blind specimen by another laboratory. Although these are not mandated practices, some clinical laboratories incorporate these practices into their QA/QC programs (Grannis et al., 1972; Gambino, 1990; Engebretson et al., 1992).

Although blind proficiency testing may be a better tool than open proficiency testing in assessing laboratory performance, as well as in obtaining a better estimate of laboratory error rates, its problems have precluded adoption as a national mandatory practice. CLIA '67 and CLIA '88 do not mandate blind PT; however, laboratories wishing to be licensed must participate in an external open PT program that is available from 14 CLIA approved proficiency test providers (incidentally, only 4 provide the full spectrum of PT). If a sanctioned, national clinical blind PT program were to be administered, the few approved proficiency test providers would have to substantially modify their current programs.

I. Error Rates in Laboratories

The issue of determining a laboratory's "error rate" has been a topic of debate and confusion in the clinical and forensic communities. This debate has been spurred on by the suggestion that a quantitative error rate of a laboratory should be used in the weighing of evidence in litigation. In the past, the term "error rate" has been used loosely, and due to misconceptions on the part of some authors and commentators, proficiency testing was once thought to measure the "true" proficiency of a laboratory (HCFA, 1990; Grannis et al., 1972). In fact, the National Research Council recommended in its first report DNA Technology in Forensic Science, that "laboratory error rates should be measured with appropriate proficiency tests (NRC, 1992)." Furthermore, the U.S. Supreme Court has offered the observation in its landmark Daubert decision that courts "should consider the known or potential rate of error" (cited in Daubert v. Merrell Dow Pharmaceuticals, Inc (see Case Citations). As a result, defendants and some legal commentators have proposed that the risk that a laboratory and/or handling errors that might falsely incriminate a suspect could be estimated from proficiency test results, and that laboratory error rate values should be combined with the random-match probability to determine one summary statistic for use in litigation (NRC, 1992; Lempert, 1991; Saks and Koehler, 1991, Koehler, 1993; Federal Judicial Center, 1994; Thompson, 1995). Currently, at least one court has held that a laboratory's proficiency testing record must accompany its estimate of the probability of a matching profile (e.g., U.S. v. Porter, see Case Citations).

The idea of modifying the random match probability is an alternative to be considered; however, the problem with this approach is that a laboratory's error rate may not be reliably determined by PTs, open or blind. As discussed earlier, the extent to which proficiency test results reflect laboratory performance depends on a number of factors, including:

(a) the similarity between PT samples and patient or crime scene specimens;
(b) the qualifications and experience level of personnel who analyze the samples;
(c) the need for a sufficient number of challenges over time in order to distinguish between an occasional random error and repeated errors within a laboratory;
(d) the fact that many errors occur in the pre-analytic and post-analytic phase while open PT measures analytic errors;
(e) the belief among proficiency test experts that with limited test data, interlaboratory PT is effective in detecting bias but fails to detect reproducibility.

(Grannis et al., 1972; Kazmierczak and Catrou, 1993; Laessig et al., 1989; Nutting et al., 1996).

The experience of commercial laboratories participating in the California Association of Crime Laboratory Directors (CACLD) DNA PT studies in 1987-1988 (Kuo, 1988) is a good example of how proficiency test results may be employed as error rates by the courts. In this proficiency trial for three commercial firms, sample handling was attributed to the two false

match errors reported by one firm involving 125 samples. This commercial laboratory was assessed an error rate that was used in courts based on single match errors.

The fundamental concern is that proficiency testing, to date, has been so limited that it is not representative of a laboratory's general performance and therefore should not be portrayed as an "error rate." Unless proficiency tests are offered with sufficient frequency over an extended time period, it may not be possible to use such results to approximate an error rate. Additionally, the CACLD results were based only on aggregate results of the proficiency tests and did not distinguish between error that occurred in the analytical and nonanalytical phases of the testing process. In order to gain a better understanding of the complete process, it is necessary to look at all errors from the beginning to the end of the testing process. The testing process consists of preanalytical, analytical, and postanalytical phases in which errors can occur. As an example, a clerical error would be considered a nonanalytical error while a technical error would be considered an analytical error.

The actual rate of laboratory error, and its preanalytical, analytical, and postanalytical components, have not been extensively studied in the setting of the current clinical laboratory (Kazmierczak and Catrou, 1993); however, previous reports on rates of laboratory error showed that the majority of mistakes were due to technical errors associated with manual portions of laboratory tests, sample switching, clerical errors due to mistakes in the manual calculation of test results and transcription errors. Advances in technology, such as automated sample processing, bar coding, standardization, quality control, result reporting, and implementation of sophisticated QA/QC programs have undoubtedly helped to reduce error rates.

Several studies have examined the origin of proficiency test errors in the clinical laboratory setting. Hoeltge and Duckworth (1987) have reported on the reasons for 583 PT failures found during the 1986 CAP Survey challenges from 27 laboratories. Klee and Forsman (1988) have discussed reasons for 89 PT failures observed within the Mayo Clinic during 1985 and 1986. Steindel et al., (1996) looked at the rate of PT failures and reasons for failure in chemistry and blood gas analysis from 670,489 challenges performed in 665 laboratories. In addition, they surveyed similar studies and categorized failure types into six major groups (methodologic, technical, clerical, survey, unexplained, or other) and into subgroups.

Findings from these studies, however, have not been consistent. Hoeltge and Duckworth, and Steindel et al. found that over 50% of PT errors were methodologic and technical. Klee and Forsman found that 28% of the total PT errors were analytic. In addition, they attributed 1/3 of the failures to the proficiency test material (e.g., matrix effect and/or degradation of PT specimen), whereas two other studies observed a much lower frequency in this area. Hoeltge and Duckworth found that 27% of errors committed were clerical errors; Steindel et al., and Klee and Forsman observed 12% and 16% in clerical errors, respectively. Steindel et al. concluded that individual analyte failure is a common event in the participating laboratories, but failures in successive or alternate events are rare. Hoeltge and Duckworth concluded that proficiency testing, generally, was a good method for determining laboratory performance while Klee and Forsman (1988) concluded that the proficiency testing process was <u>not</u> a good indicator of laboratory analytic performance due to inadequate statistical power of proficiency testing.

Currently, the forensic DNA community has underscored the limitations of assessing laboratory error rates through the use of proficiency testing. The National Research Council (NRC) has changed its position on the role of proficiency testing. The 1996 report declared that "...proficiency testing is to improve laboratory performance by identifying problems that need to

be corrected" and "...is not designed to measure error rates (NRC, 1996)." In addition to the above points, the NRC notes that when an error is discovered, it is typically investigated so thoroughly that corrective and preventive actions are taken. Therefore, the laboratory is not likely to make the same error again, so the error probability is correspondingly reduced (NRC, 1996). Furthermore, the committee members of the 1996 NRC report, believed that a "calculation that combines error rates with match probabilities is inappropriate." The committee brought to light several important points in which assessing laboratory error rates would be problematic (NRC, 1996). Those concerns are as follows:

(a) "The general error rate of a laboratory over time is not the concern, but whether the laboratory doing the DNA testing in a particular case made a critical error. There are too many variables and inadequate methods to determine this;

(b) To estimate accurately, from proficiency test results, the overall rate at which a laboratory declares nonmatching samples to a match, as has been suggested, would require a laboratory to undergo an unrealistically large number of proficiency trials;

(c) It has been suggested that pooling of proficiency test results across laboratories could produce an 'industry wide' error rate. However, better labs could be penalized and multiple errors on a single test by one lab could affect the overall estimated false-match error rate;

(d) The commercial laboratory, participating in the CACLD PT program, made errors that occurred in the first two years of its operation. Since then, the laboratory has not committed any additional proficiency test error, therefore, it would be inappropriate to use the original error rate."

The committee reported that the risk of error should be "properly considered case by case, taking into account the record of the laboratory performing the tests, the extent of redundancy, and the overall quality of the test results" (NRC, 1996).

So, although the consensus of the field appears to be that error rates cannot be reliably determined through proficiency testing, proficiency test results can be useful in comparing performance among laboratories. Furthermore, there is also consensus that the proficiency test data can be used to monitor a laboratory's performance over time.

J. Proficiency Testing in DNA Identification Laboratories

1. Background

As in clinical laboratories, the importance of testing the validity and reliability of scientific test results in the crime laboratory is paramount. Until the mid 1970's, however, there were virtually no procedures in place for this testing. In 1974, a grant from the National Institute of Law Enforcement and Criminal Justice (LEAA) to the Forensic Sciences Foundation (FSF) enabled the FSF to conduct research on developing a proficiency testing program for crime laboratories. This study revealed serious problems in the examination and interpretation of results from some types of specimens (Peterson et al., 1978). The study concluded there was a need for the commitment of greater resources to these laboratories, along with improved education and training opportunities, implementation of accreditation and certification programs, as well as the need for ongoing proficiency testing and quality assurance programs (Peterson and Markham, 1995).

The British Forensic Science Service (FSS) has long been cited for its demanding quality assurance standards, including proficiency testing which began in 1969. Margaret Pereira has

described FSS's quality assurance program, noting it included both open and blind (undeclared) trials which enter laboratories disguised as genuine cases (Pereira, 1985). Although blind trials are much more difficult to construct, she comments their advantage is that they test "the whole system"; from receipt of evidence and quality of scientific work to the laboratory report, as well as the time required to complete the case. David Werrett of the Home Office Central Research Establishment noted at the Banbury Conference in 1988 that the British had already established undeclared DNA trials (Werrett, 1988).

Progress has been notable on many of these fronts over the past twenty years, including the continuation and expansion of proficiency testing at local, state, federal, and international laboratories. In 1987-1988, the California Association of Crime Laboratory Directors organized proficiency trials for three commercial facilities using simulated DNA evidence samples (Kuo, 1988). The American Association of Blood Banking (AABB) also started a DNA proficiency testing program by adding a DNA module to their 1991 Parentage Specimen Program (PSP) which was subsequently replaced by the AABB/CAP Parentage Identity (PI) survey. Collaborative Testing Services (CTS), in conjunction with the American Society of Crime Laboratory Directors and the Forensic Sciences Foundation, began one of the first DNA PT programs by adding a DNA module to its physiological fluids offerings within its forensic laboratory testing program in 1991. In 1993, the College of American Pathologists began proficiency testing for both the forensic and parentage laboratories under its forensic identification survey, beginning with 41 participants and growing to 80 within a year (Reeder, 1995). In addition to CTS and CAP, DNA proficiency testing trials are also available from Serological Research Institute (SERI); and Cellmark Diagnostic's International Quality Assurance Survey (IQAS).

The Bureau of Justice Statistics published survey results in February 2000 based on replies from 108 forensic laboratories in the United States performing DNA testing (Steadman, 2000). The survey attempted to establish "baseline information" on all publicly funded forensic laboratories performing DNA testing in 1997 and 1998. All laboratories required their analysts to undergo proficiency testing every 180 days or less, and to follow TWGDAM guidelines. While almost 90% of labs required technical leaders to complete these proficiency tests, one-third required technicians to do so. While some laboratories only required proficiency tests once a year, most required the tests every six months.

2. Proficiency Testing Standards

The ASCLD Laboratory Accreditation Board (ASCLD-LAB) requires laboratories seeking (voluntary) accreditation to establish and maintain a "quality system", appropriate for its casework. Proficiency testing is cited as an "integral component" of QA programs and requires laboratories to subscribe to an external proficiency test provider in all disciplines in which they seek accreditation. Non DNA examiners are to complete at least one PT test annually in their functional area and DNA examiners are required to complete two PT tests annually, one of which must be from an approved external test provider (American Society of Crime Laboratory Directors, 1994).

Recognizing the increasing use and importance of proficiency testing in quality assurance, the TWGDAM "Guidelines for a Quality Assurance Program for DNA Analysis" devotes an entire section to proficiency testing. Section 9 of the Guidelines states, "Participation in a proficiency testing program is a critical element of a successful QA program and is an essential requirement for any laboratory performing forensic DNA analysis" (TWGDAM, 1995)

TWGDAM discusses both open and blind proficiency testing, noting it is "highly desirable" for the DNA laboratory to participate in a blind proficiency testing program that "realistically simulates" actual casework in order to evaluate "all aspects of the laboratory examination procedure" (TWGDAM, 1995). Laboratories implementing such blind programs should be tested in this fashion at least once a year. Laboratory quality assurance coordinators have the responsibility of documenting and reviewing PT results, noting any discrepancies and taking appropriate corrective action.

Because of the increasing number of quality assurance programs requiring proficiency testing, the TWGDAM Quality Assurance Subcommittee joined with the DNA Proficiency Review Committee (PRC) of ASCLD-LAB to produce Guidelines for DNA Proficiency Test Manufacturing and Reporting (1994) to set standards for commercial providers of DNA test samples. These guidelines also set standards for the personnel, facilities, and procedures used by the manufacturers, as well as quality control procedures they must follow in producing PT specimens. The ASCLD-LAB DNA Proficiency Review Committee (PRC) performs on-site visits of proficiency test manufacturers and providers to ensure compliance with these guidelines. Recently, ASCLD announced those PT providers which have been approved to service ASCLD-LAB accredited laboratories. In addition, The American Board of Criminalists (ABC) certifies individuals based upon their educational background, experience, and performance on a written examination. Those seeking Fellow status in a specialty area like DNA testing must also supply proficiency test performance results, and must submit acceptable PT results annually thereafter to maintain their certification.

3. Evaluation of Proficiency Testing

Several study groups have had the opportunity to evaluate the quality of parentage and forensic DNA testing. The Office of Technology Assessment (OTA) of the U.S. Congress was asked to investigate the propriety of forensic DNA testing. The OTA issued a report in 1990 which strongly endorsed the types of DNA testing which were being used in forensic laboratories and declared, "The Office of Technology Assessment found that the forensic uses of DNA tests are both reliable and valid when properly performed and analyzed by skilled personnel" (OTA, 1990).

The National Research Council (NRC) issued its report in 1992 entitled DNA Technology in Forensic Science (NRC, 1992). The NRC agreed with the OTA report that forensic uses of DNA tests are reliable and valid when properly performed. With regard to proficiency testing, the NRC endorsed the TWGDAM guidelines for quality assurance, which is to say they recommended regularly scheduled proficiency testing as a way of measuring laboratory error rates (false positives and false negatives). Like the 1990 OTA Report, the NRC also noted on the error rates of the commercial laboratories in the CACLD proficiency tests in 1978-1988. The NRC report recommended that error rates be "continually estimated in blind proficiency testing" (NRC, 1992). The Committee went on to advise that errors occur in the best laboratories and that "error rates" need continuous review and adjustment: "One purpose of regular proficiency testing under standard case conditions is to evaluate whether and how labs have *taken corrective action to reduce error rates*" (NRC, 1992). They noted such tests would ideally involve blind tests of representative case materials. In addition, the report stated that regulation by a government agency is necessary to oversee the voluntary accreditation programs of professional organizations. They recommended that the Department of Health and Human Services in conjunction with the Department of Justice be legislatively mandated to regulate forensic

laboratories. With regards to estimating error rated with proficiency tests, the NRC changed its stand in 1996 and in a report entitled The Evaluation of Forensic DNA Evidence (NRC, 1996) stated, proficiency testing "...is not designed to measure error rates," however, proficiency testing "...is one of the best ways of ensuring standards and....should be used to improve laboratory performance by identifying problems that need to be corrected." The NRC also recommended that "laboratories should participate regularly in proficiency tests, and the results should be available for court proceedings."

4. The DNA Identification Act of 1994

A joint hearing on forensic DNA analysis before the Subcommittee on Civil and Constitutional Rights in 1991 had influenced Congress and in the language contained in the DNA Identification Act passed in 1994 (P.L. 103-322, 1994). Testimony was offered that there was a need for external blind proficiency testing based on the assertion that clinical laboratories are mandated by CLIA to participate in blind proficiency testing (US Congress, Hearings, No. 30, 1991). CLIA, however, does not have such a requirement, nor does such a national program exist. Additional testimony was offered including: creating an independent board of scientists to set standards; licensing of forensic laboratories and their personnel; and privacy issues surrounding DNA type data basing.

The DNA Identification Act of 1994 (P.L. 103-322, 1994) establishes a federal framework for setting standards on quality assurance and proficiency testing. The law created the DNA Advisory Board (DAB) of which the members are appointed by the FBI Director. The DAB consists primarily of forensic scientists, molecular geneticists, population geneticists, and members knowledgeable in law and ethics. The law states, "The advisory board shall develop, and if appropriate, periodically revise, recommended standards for quality assurance, including standards for testing the proficiency of forensic laboratories, and forensic analysts, in conducting analyses of DNA." In fact, the board has been developing a set of standards for DNA testing.

The DNA Act also authorizes the appropriation of grant funding to laboratories which establish DNA databases and repositories and which provide DNA testing services. One contingency for receiving funding is the implementation of a strict QA program adhering to the DAB's standards for DNA testing. Within those standards includes participation in a regularly-scheduled, external PT program. The DNA Act also directs the National Institute of Justice (NIJ) to report to Congress that "(1) a blind external DNA proficiency testing program already exists; or (2) blind external DNA proficiency testing is not feasible; (3) or NIJ has entered into contract with, or made a grant to, an appropriate entity for establishing, or has taken other appropriate action to ensure that there is established, not later than two years after the date of enactment of the Act, a blind external proficiency testing program for DNA analyses, which shall be available to the public and private laboratories (P.L. 103-322, 1994)."

5. Summary of Recent Commercial Proficiency Testing

Currently, DNA proficiency testing trials are available from Collaborative Testing Services (CTS), the College of American Pathologists (CAP), the Serological Research Institute (SERI), and Cellmark Diagnostic's International Quality Assurance Survey (IQAS). In general, each sample pack consists of bloodstains and/or semen stains of which there is a crime scene stain and a combination of suspect and victim stains. The object of these proficiency tests is to correctly exclude or include suspect/victim stains from the crime scene stain. Each of these DNA PT programs allows its participants to report information pertaining to methodology, band

sizing data from RFLP analysis, and discrete results from PCR-based DNA testing. Per TWGDAM guidelines, each of the providers manufacture and distribute two programs per year.

(a) Collaborative Testing Services

Laboratories have performed quite well in CTS proficiency tests over the past five years, with the 93-C DNA Profiling Report reporting that all laboratories reporting conclusions were correct, with corresponding "remarkable" consistency of their RFLP band size data. All substantive responses from laboratories using PCR were correct, although there was a single incorrectly reported allele by one laboratory (CTS, 1993). From 1995 to 1997, 2140 tests were distributed and a total of 509 participants reported RFLP procedures and 863 participants reported PCR procedures (CTS, 1995a, 1995b, 1996a, 1996b, 1997a, 1997b). Of these, there were three incorrect conclusions due to clerical/reporting errors. In one test, there were eight failures to exclude a suspect as a possible source of the crime scene bloodstain; however, these results were not the result of errors but because these laboratories analyzed the samples only for the DQA1 loci which did not adequately discriminate the samples. In general, RFLP bandsizing results were remarkably consistent with the group mean. In fact, less than 1% of all band sizes reported were more than +/-2.5% from the group mean. A total of 56 (6.5%) PCR results were not consistent with the consensus result. These inconsistencies were due to incorrect identification of alleles, clerical errors, and sizing inconsistencies. Still, none of the 56 inconsistent PCR results led the laboratories to submit an incorrect inclusion or exclusion.

In 1998 and 1999, laboratories performed very well on three additional CTS proficiency tests (CTS, 1998a, 1998b, 1999). In 1998, on two separate tests, more than 99% of the reported conclusions were correct. The three laboratories incorrectly including the female item as the source of the bloodstain in 9815 reported only D1S80 results, and therefore could not discriminate the source of the stain. In 1999, a mixture was used as a part of the test which increased its difficulty. In 99-511, most of the incorrect reports were false exclusions, plus the CTS manufacturers also noted that some of the difficulties with this test may be the result of how the sample mixture was prepared. It also became clear that a greater proportion of laboratories were beginning to use STR's as their preferred technique, and RFLP was being used less after.

(b) College of American Pathologists

All laboratories responding to the 1993 pilot survey analyzed samples correctly. In the 1994 survey results, laboratory performance was "outstanding" with DNA testing "more accurate and precise than most analytical tests run in clinical laboratories (Reeder, 1995)." RFLP results from the latest CAP Forensic Identity proficiency testing programs (1996 FID-A and FID-C) indicate that, "...the range of band values among respondents continue to narrow," and "...the variation of reported results among participating laboratories demonstrated a high degree of accuracy and precision (CAP, 1996a, 1996b)." In the 1996 FID-A program, 41 of 1096 (3.7%) participating laboratories reported PCR results that were inconsistent with the reference laboratory; however, only one laboratory out of ninety-three reported an incorrect conclusion which was attributed to sample switching. In the 1996 FID-C program, 51 out of 1246 (4.1%) reported PCR results that were inconsistent and only one laboratory out of ninety-three reported an incorrect conclusion. Finally, in the 1997 FID-A program two respondents out of ninety-six generated inconsistent results on two specimens leading to incorrect conclusions (CAP, 1997).

(c) Cellmark Diagnostics

A total of four tests were distributed by Cellmark Diagnostic's IQAS program during 1996 and 1997 (IQAS, 1996a, 1996b, 1997a, 1997b). A total of 466 sample sets were distributed and 382 analyses were reported. Of all the sample sets tested by laboratories, every laboratory made correct identifications. However, no additional data are available regarding the accuracy of laboratories reporting results.

(d) The Spanish and Portuguese Working Group (GEP) of the International Society for Forensic Genetics (ISFG) (Gómez and Carracedo, 2000)

The Spanish and Portuguese Working Group of the International Society for Forensic Genetics (GEP-ISFG) is mostly composed of forensic laboratories in Spain, Portugal, France and from the Portuguese and Spanish speaking countries in America. In 1992, the GEP-ISFG started issuing collaborative exercises on DNA profiling, and began a proficiency testing programme in 1995. The number of participating labs has increased annually, with 46 laboratories enrolled in 1999. They found many of the same changes in methods as CTS has discovered; while the percentage of labs using SLP methodologies has decreased, the fraction using PCR/STRs has increased. This also coincides with greater uniformity in testing results. The results of PCR-based analyses showed that only a few mistakes were found, and these errors were primarily due to poor quality ladders or techniques, transcription errors, and lack of detection of intermediate alleles. Meanwhile, errors in some statistical programs were also found. Overall, despite the differences in methodologies and the difficulties of the cases, the results were deemed "satisfactory" by the organizers.

6. Conclusion

The NIST in concert with the FBI and TWGDAM investigated the performance of the forensic DNA testing community on the basis of proficiency testing (NRC, 1996). The conclusions drawn from the data include: 1) proficiency testing provides an invaluable service to the parentage and forensic DNA testing communities; 2) the parentage and forensic DNA testing communities, with noted exception, are performing consistently and reliably; 3) inter-laboratory RFLP results are very accurate, but exhibit a variable degree of precision; and 4) intra-laboratory RFLP results are highly reproducible (Reeder, 1995).

K. Conclusion

With the passage of the DNA Identification Act of 1994, the forensic DNA community has renewed its interest in investigating the quality of laboratory results and in determining the role of blind proficiency testing in forensic DNA laboratories. This parallels the clinical laboratory community's experience where adverse publicity claiming high error rates led to the passage of the Clinical Laboratory Improvement Acts (CLIA) of 1967 and 1988. Subsequently, the federal government mandated proficiency testing, with few exceptions, for all clinical laboratories.

Proficiency testing is used in laboratories to serve as a mechanism for self-improvement and to assess quality performance. The primary objective of proficiency testing is to ensure the quality of a laboratory through critical self-evaluation and self-education. The process involves interlaboratory comparisons of PT data and/or identifying problems that contribute toward error within the laboratory. Subsequently, these problems are corrected so that may be avoided in the future. The end result is better laboratory quality, and increased precision and accuracy of test results.

The other purpose of proficiency testing is that is to serve as a method for assessing laboratory performance. In fact, proficiency testing performance has been correlated with various factors in the laboratory. Laboratory characteristics including type, size, test volume, workload, and accreditation status have been related to performance. Personnel standards such as qualifications, education, and training have also been related to performance. Numerous studies have also shown consistent and statistically significant improvement in the performance of laboratories on proficiency tests over time. This may be due to advances in quality control procedures, testing methodologies, and automation or simply the fact heightened vigilance associated with the PT program itself has caused laboratories to improve its quality in testing.

One of the limitations of open PT, however, is because it centers on the analytic stage of the testing process, it does not capture errors that may occur outside of analytic procedures. Commentators have expressed that proficiency test results are indicative of the best a laboratory can produce. This is because open PTs are recognized as such and special treatment may be given to those samples. To remedy the limitations of open proficiency tests, it has been suggested that blind PT would be a "truer" measure of laboratory performance. In fact, the TWGDAM "Guidelines for a QA Program for DNA Analysis" also note it is "highly desirable" for the DNA laboratory to participate in a blind proficiency testing program annually that "realistically simulates" actual casework. In many comparative studies, assessments of qualitative results consistently showed that blind PT samples resulted in lower correct response rates and, in particular, higher rates of false negatives, and that improvements in blind proficiency tests generally lagged over time (when compared with open testing). A study by Glenn and Hathaway (1979) found the analytical results from the blind specimens to be comparable to the open samples, but that the blind process was particularly useful in detecting problems in the pre- and post analytical phases.

Although blind trials of laboratories have unquestionably provided important insights, the complexity and expense limits blind PT as an external QA tool. The logistics and resource requirements of conducting blind testing on a national scale for hundreds or thousands of laboratories are formidable. Forensic laboratories create the added complexity of requiring law enforcement agencies to create a fictitious criminal case investigation with associated case reports and chain of custody (paperwork) requirements. Such cases still have to be introduced through a series of supervisors and analysts without being detected, and the simulated evidence examined in a reasonable time period, so that results may be compiled by the central issuing organization. Efforts to speed up the analysis of evidence by introducing a particularly serious case or having investigators apply pressure to the analysts may also compromise the test. Many authors conclude that blind testing can be successful, but only on a small scale, involving a single laboratory system or group of laboratories involving testing of manufactured specimens or the re-testing of actual samples. Given the foregoing problems and issues, neither CLIA '67 nor CLIA '88 mandate blind proficiency testing of clinical laboratories.

There are still many conflicting opinions on the uses, limitations, and advantages of proficiency testing; however, the clinical laboratory profession has recognized the importance of proficiency testing in quality assurance. Instead of dwelling on the limitations of proficiency testing, clinicians are focusing on the potential contributions of proficiency testing and strengthening the process by finding better methods to improve laboratory quality, precision, and accuracy.

L. References

AABB. Standards for Parentage Testing Laboratories, American Association of Blood Banks, Bethesda, MD, 1991.

Anderson TD, Roby RK, Weedn VW. Analysis of the Data from the College of American Pathologists - Forensic and Parentage Identity DNA Proficiency Testing Programs. Armed Forces Institute of Pathology. Rockville, MD. 1995.

ASCLD. Laboratory Accreditation Board Manual. American Society of Crime Laboratory Directors-Laboratory Accreditation Board. January 1994. pp. 17-21.

ASHI. ASHI Procedure Manual: Standards for Histocompatibility Testing, American Society of Histocompatibility and Immunology. Lenexa, KS. 1995.

Ashton P. Testimony of the American Society for Cytotechnology before U.S. Congress. Senate Committee on Human Resources, Cancer detection in women. Hearing , January 18, 1988. No. 100-259, pp 167-224.

AuBuchon JP. Mathematical Model for Comparing Proficiency Testing Results Across Analytical Methods. Clin Chem 1991;37:265-269.

Bartola J. The Importance of an Effective Proficiency Testing Program to the Regulation of Clinical Laboratories: The View from One State. Arch Pathol Lab Med 1988;112:368-370.

Belk WP, Sunderman FW. A survey of chemical analyses in clinical laboratories. Am J Clin Pathol 1947;17:853-61.

Belsey R, Baer DM. Proficiency of office microbiology testing. Clin Lab Med. 1986;6:345-354.

Bloch MJ. Cembrowski GS, Lembesis GJ. Longitudinal Study of Error Prevalence in Pennsylvania Physicians' Office Laboratories. JAMA 1988;260/2:230-235.

Boeckx RL. Government Regulation: Can It Guaranty Quality in the Clinical Laboratory? Discussion Session I and Discussion Session II. Clin Chem 1992;38/7:1218-1250.

Boone DJ. Evaluating Laboratory Performance: Historical and Governmental Perspectives. Arch Pathol Lab Med 1988;112:354-356.

Boone DJ, Hearn T, Lewis S. Assessment of extent to which laboratories compare results before reporting in national laboratory performance surveys, abstracted. Clin Chem 1985;31:115.

Boone JD et al. Laboratory Evaluation and Assistance Efforts: Mailed, On-Site and Blind PT Surveys Conducted by the Centers for Disease Control. AJPH 1982;72:1364-1368.

Boulton FE et al. Guidelines for compatibility testing in hospital blood banks. Clin. Lab. Haemat. 1987; 9:333-341.

International Quality Assessment Scheme. Forensic Analysis: Distribution 9603. Cellmark Diagnostics. Germantown, MD, 1996a.

International Quality Assessment Scheme. Forensic Analysis: Distribution 9604. Cellmark Diagnostics. Germantown, MD, 1996b.

International Quality Assessment Scheme. Forensic Analysis: Distribution 9701. Cellmark Diagnostics. Germantown, MD, 1997a.

International Quality Assessment Scheme. Forensic Analysis: Distribution 9702. Cellmark Diagnostics. Germantown, MD, 1997b.

Cembrowski GS, Vanderlinde RE. Survey of Special Practices Associated With College of American Pathologists Proficiency Testing in the Commonwealth of Pennsylvania. Arch Pathol Lab Med 1988;112:374-376.

Collaborative Testing Services. DNA Profiling: Report No. 93Q. CTS, Inc. Herndon, VA, 1993.

Collaborative Testing Services. DNA Profiling & Physiological Fluids Analysis: Report No. 9503. CTS, Inc. Herndon, VA, 1995a.

Collaborative Testing Services. DNA Profiling & Physiological Fluids Analysis: Report No. 9515. CTS, Inc. Herndon, VA, 1995b.

Collaborative Testing Services. Serology/DNA Profiling: Report No. 9603. CTS, Inc. Herndon, VA, 1996a.

Collaborative Testing Services. Serology/DNA Profiling: Report No. 9615. CTS, Inc. Herndon, VA, 1996b.

Collaborative Testing Services. DNA Profiling: Report No. 9703. CTS, Inc. Herndon, VA, 1997a.

Collaborative Testing Services. DNA Profiling: Report No. 9715. CTS, Inc. Herndon, VA, 1997b.

College of American Pathologists, Forensic Identity Program: 1996 Set FID-A Critique, CAP, Northfield, IL, 1996a.

College of American Pathologists, Forensic Identity Program: 1996 Set FID-C Critique, CAP, Northfield, IL, 1996b.

College of American Pathologists, Forensic Identity Program: 1997 Set FID-A Critique, CAP, Northfield, IL, 1997a.

Copeland BE, Rosenbaum JM. Organization, planning, and results of the Massachusetts society of Pathologists Regional Quality Control Program. Am J Clin Pathol 1972;57:676-688.

Daly JC, Asmus KE. Laboratory performance in proficiency testing: The New York State experience. Environ. Sci. Technol 1985;19:8-13.

Damato J et al. The Department of Army Quality Assurance Program for Human Immunodeficiency Virus Antibody Testing. Laboratory Medicine 1988;19:577-580.

Davis KH, Hawks RL, Blanke, RV. Assessment of Laboratory Quality in Urine Drug Testing: A Proficiency Testing Pilot Study. JAMA 1988;260:1749-1754.

Dept. of Health, Education, and Welfare, Public Health Service, Centers for Disease Control; Clinical Laboratory Improvement Act of 1967. Part F. Title III, Public Health Service Act. Section 353, 1967.

Dewald GW et al. Pilot studies for Proficiency Testing Using Fluorescence In Situ Hybridization with Chromosome-specific DNA Probes. Arch PAthol Lab Med 1997;121:359-367.

Dewald GW et al. Toward Quality Assurance for Metaphase FISH: A Multicenter Experience. Am J Med Gen 1996;65:190-196.

Duckworth JK. Proficiency Testing: Its Role in a Voluntary Clinical Laboratory Accreditation Program. Arch Pathol Lab Med 1988;112:346-348.

Edinger SE. The Medicare and Clinical Improvement Act of 1967 Proficiency Testing Requirements and Its Relationship to the Private Sector. Arch Pathol Lab Med 1988;112:357-362.

Ehrmeyer SS, Laessig RH. Effect of Legislation (CLIA'88) on Setting Quality Specifications for US Laboratories. Scand J Clin Lab Invest 1999; 59: 563-568.

Elevitch P. Ed. Data ReCAP 1970-1980, Northfield, III: College of American Pathologists; 1981.

Engebretson MJ, Cembrowski GS. Achieving the Health Care Financing Administration Limits by Quality Improvement and Quality Control. Arch Pathol Lab Med 1992;116:781-787.

Erhmeyer SS, Laessig RH. An evaluation of the ability of proficiency testing programs to determine intralaboratory performance: Peer group statistics vs. Clinical usefulness limits. Arch Pathol Lab Med 1988;112:444-448.

Erickson R, Driscoll C, Dvorak L, et al. Factors related to accuracy in office cholesterol testing. J Fam Pract. 1991;33:457-461.

Esche CA, Groff JH. ELPAT Program Report: Background and Current Status (October 1996). AIHA Journal 1997;58:59-63.

Estevez EG, A Program for In-House Proficiency Testing in Clinical Microbiology. Am J Med Tech 1980;46:102-105.

Federal Judicial Center. Reference Manual on Scientific Evidence. Washington, DC. 1994.

Freund, Journal of Quality Technology, 1985.

Frings CS, White RM, and Battaglia DJ. Status of Drugs-of-Abuse Testing in Urine under Blind Conditions: An AACC Study. Clin Chem 1989;35/5:891-894.

Gambino R, Mallon P, Woodrow G. Managing for Total Quality in a Large Laboratory. Arch Pathol Lab Med. 1990;114:1145-1148.

Garber CC, Witte DL. Quality for Tomorrow: By Design or By Checking? Clinical Chemistry 1997; 43: 864-865.

Gilbert RK. Accuracy of clinical laboratories studied by comparison with definitive methods. Am J Clin Pathol 1978;70:450-470.

Glenn GC. College of American Pathologists Conference XIII on the Evaluation of Proficiency Testing Results for Quantitative Methods in Relation to Clinical Usefulness: Concluding Remarks. Arch Pathol Lab Med 1988;112:471-472.

Glenn GC, Hathaway TK. Quality Control by Blind Sample Analysis. Am J Clin Pathol 1979;72:156-162.

Gómez J, Carracedo A. The 1998-1999 Collaborative Exercises and Proficiency Testing Program on DNA Typing of the Spanish and Portuguese Working Group of the International Society for Forensic Genetics (GEP-ISFG). Forensic Science International 2000; 114: 21-30.

Grannis GF, Grumer HD, Lott JA, Edison JA, McCabe WC. Proficiency Evaluation of Clinical Laboratories. Clin Chem 1972;18:222-236.

Grannis GF. Studies of the reliability of constituent target values established in large interlaboratory survey. Clin Chem 1976;22:1027-1036.

Growe GH. Standards, Accreditation and Proficiency Testing in the Management of Blood and Blood Products in Hospitals. Can Med Assoc J 1996;154:755.

Hamlin W. Proficiency Testing as a Regulatory Device: A CAP Perspective. Clinical Chem 1992;38/ 7:1234-1238.

Hammond HC. Federal Regulation of Clinical Laboratories: A Prospective Public Policy Analysis. Arch Pathol Lab Med 1988;112:363-367.

Hammond HC. Federal Regulation of Clinical Laboratories: A Prospective Public Policy Analysis. Arch Pathol Lab Med 1988;112:363-367.

Hansell JR, Haven GT. Changes in level of precision of common ligand assays during a seven year interval. Amer J Clin Pathol 1979:72:320-325.

Hansen HJ, Caudill SP, Boone JD. Crisis in Drug Testing: Results of a CDC Blind Study. JAMA 1985;253:2382-2387.

Hartmann AE, Naito HK, Burnett RW, et al. Accuracy of participant results utilized as target values in the CAP chemistry survey program. Arch Pathol Lab Med 1985;109:894-903.

Health Care Financing Administration. Medicare, Medicaid and CLIA Programs: Revision of laboratory regulations; final rule with request for comments. Fed Reg 1990a;55:9534-610.

Health Care Financing Administration. Medicare and Laboratory Certification Programs; enforcement procedures for laboratories. Fed Reg 1991;56:13430-45.

Health Care Financing Administration. State Operations Manual: Provider Certification. Washington, DC, Health Care Financing Administration, 1986.

Health Care Financing Administration. 42 CFR, Parts 405, 416, 440, 482, 483, 488, and 493: Medicare, Medicaid and CLIA programs; regulations implementing the Clinical Laboratory Improvement Amendments of 1988; proposed rule. Fed Reg 1990b;55: 20896-959.

Hinkel C. Regulation of Proficiency Testing Under the March 14 Rule. Clinical Chemistry 1992;38/7:1251-1253.

Hodnett J. Proficiency Testing: We All Do It — But What Do the Results Mean? Laboratory Medicine 1999; 30: 316-323.

Jenny RW, Jackson-Tarentino KY. Causes of Unsatisfactory Performance in Proficiency Testing. Clinical Chemistry 2000; 46: 89-99.

Jonakait, RN. Forensic Science: The Need for Regulation. Harvard Journal of Law and Technology 1991;4:109-191

Jones, RN, Edson DC. Microbiology Resource Committee of CAP. Antimicrobial susceptibility testing trends and accuracy in the United States. Arch Pathol Lab Med. 1991:115;429-436.

Kaye DH, Sensabaugh GF, Jr. Reference Guide on DNA Evidence in Reference Manual on Scientific Evidence, second edition. Federal Judicial Center, 2000.

Kazmierczak SC, Catrou PG. Laboratory Error Undetectable by Customary Quality Control Quality Assurance Monitors. Arch Pathol Lab Med 1993;117:714-718.

Key PE, Patey AL, Rowling S, Wilbourn A, Worner FM. International Proficiency Testing of Analytical Laboratories for Foods and Feeds from 1990 to 1996: The Experiences of the United Kingdom Food Analysis Performance Assessment. J. AOAC Intl 1997;80:895-899.

Kirby LT. DNA Fingerprinting: An Introduction, New York: Stockton Press, 1990.

Klee GG, Forsman RW. A User's Classification of Problems Identified by Proficiency Testing Surveys. Arch Pathol Lab Med 1988;112:371-373.

Koehler JJ. DNA matches and statistics: important questions, surprising answers. Judicature 1993;76(5):222-229.

Kuo M (Chair). California Association of Crime Laboratory Directors: DNA Committee Report to the Board of Directors, Report #6, Orange County Sheriff Coroners Crime Laboratory. Santa Ana, CA, 1988.

Laessig et al. Limitations of Proficiency Testing under CLIA '67. Clin Chem. 1992b;38/7: 1237-1244.

Laessig RH, Ehrmeyer SS, Lanphear BJ, Burmeister BJ, Hassemer DJ. Limitations of Proficiency Testing under CLIA '67. Clin Chem 1992a;38:1237-1244.

Laessig RH, Ehrmeyer SS. Proficiency Testing Programs-Promises, Progress, and Problems. Arch Pathol Lab Med 1988;112:329-333.

Laessig et al. The Effectiveness of Proposed CDC/HCFA Criteria to Assess Intralaboratory Performance Through Proficiency Testing Programs. Clin Chem 1989;35:1088.

Lamotte LC, Guerrant GO, Lewis DS, Hall CT. Comparison of Laboratory Performance with Blind and Mail Distributed PT Samples. Pub Health Rep 1977;92:554-560.

Lanphear et al. Review of Actual Proficiency-Testing Performance under CLIA '67 March 14, 1990 Rules: Perspective from the First Year's Data. Clinical Chemistry 1992;38:1254-1259.

Lawson NS, Gilmore BF, Tholen DW. Multiprogram characterization or laboratory bias, precision, and total error proposal for improved assessment with shared external and expanded internal quality control pools. Arch Pathol Lab Med. 1988;112:454-461.

Lempert R. Some Caveats Concerning DNA as Criminal Identification Evidence: With Thanks to the Reverend Bayes. Cardozo L Rev 1991;13(3):303-341.

Macro Systems Inc. Final report on the Assessment of Clinical Laboratory Regulations. National Technical Information Service publication 862104401/AS. Springfield, Ill, US Department of Commerce, 1986.

Mason MF. Some Realities and Results of PT of Laboratories Performing Toxicological Analyses. J of Anal Toxicol 1981;5:201-208.

McKenna JA, Cecil JS, Coukos P. Reference Guide on Forensic DNA Evidence in Reference Manual on Scientific Evidence. Federal Judicial Center. 1994.

National Research Council, DNA Technology in Forensic Science, Washington, DC. National Academy Press, 1992.

National Research Council, The Evaluation of Forensic DNA Evidence, Washington, DC: National Academy Press, 1996

National Institute of Standards and Technologies. Certificate of Analysis. Standard Reference Material 2390 and 2391: DNA Profiling Standards. August, 1992.

Nutting et al. Problems in Laboratory Testing in Primary Care. JAMA 1996;275:635-639.

Office of Technology Assessment (U.S. Congress), Genetic Witness: Forensic Uses of DNA Tests (OTA-BA-438), Washington, DC: Government Printing Office (GP0052-003-01203-1), 1990.

Parvin CA. Quality-Control (QC) Performance Measures and the QC Planning Process. Clinical Chemistry 1997; 43: 602-607.

Peddecord KM et al. Analytic Results of HIV-1 Testing Using Blind Proficiency Testing. Clin Lab Sci 1992b;5:165-171.

Peddecord MK and Hammond HC. Clinical Laboratory Regulation under the Clinical Laboratory Improvement Amendments of 1988: Can It Be Done? Clin Chem 1990;36/12:2027-2035.

Peddecord KM et al. Identification of Nonanalytic Issues in HIV- 1 Antibody Testing Using Blind Proficiency Testing. Clin Lab Sci 1992c;5:172-176.

Peddecord KM, Taylor RN. Evaluation of Laboratory Quality Assurance Activities: Final Report. Washington DC. Department of Health and Human Services 1983. Grant HS04767.

Peddecord KM. A critical review of personnel standards. Clin Chem 1996;42:790-796.

Peddecord KM et al. Methods for a Model Blind Proficiency Testing System. Clin Lab Sci 1992a;5:160-164.

Pereira M. Quality Assurance in Forensic Science. Forensic Science International 1985;28:1-6

Peterson JL, Fabricant EL, Field KS, Thornton JI. Crime Laboratory Proficiency Testing Research Program. U. S. Government Printing Office, Washington, D.C., 1978.

Peterson JL, Markham P. Crime Laboratory Proficiency Testing Results, 1978-1991, Parts 1 and 2. Journal of Forensic Sciences 1995;40:1009-1029.

Phillipou G, Farrant RK, Phillips PI. Computer based quality assessment of hospital capillary blood glucose monitoring. Diabet Med. 1990;7:234-237.

Proceedings of the International Symposium on the Forensic Aspects of DNA Analysis: June 19-23, 1989, Forensic Science Research and Training Center, FBI Academy, Washington, D.C., Government Printing Office, 1989.

Public Law No. 100-578. 100th Congress. Clinical Laboratory Improvement Amendments of 1988. Stat 42 USC 201. H.R. 5471. Oct 31, 1988.

Public Law 90-174. Clinical Laboratory Improvement Act of 1967. Code of Federal Regulations, Title 42, Part 72, 1967.

Public Law No. 103-322. DNA Identification Act of 1994, Subtitle C of the Violent Crime Control and Law Enforcement Act of 1994. September 13, 1994.

Reeder D. DNA Testing Survey Takes Hold in Crime Laboratories. CAP Today. April 1995.

Rej R, Jenny RW. How Good Are Clinical Laboratories? An Assessment of Current Performance. Clin Chem 1992;38/7:1210-1217.

Richardson H, Wood D, Whitby J, Lannigan R, Fleming C. Quality Improvement of Diagnostic Microbiology Through a Peer-group Proficiency Assessment Program: A 20-year Experience in Ontario. Arch Pathol Lab Med 1996;120:445-455.

Saks MJ, Koehler JJ. What DNA Fingerprinting Can Teach the Law about the Rest of Forensic Science. Cardozo L Rev 1991;13 (2-3):361-372.

Salkin IF, Limberger RJ, Stasik D. Commentary on the Objectives and Efficacy of Proficiency Testing in Microbiology. J Clin Micro 1997;35:1921-1923.

Schalla WO, Blumer SO, Taylor RN, et al. HIV blind performance evaluation: a method for assessing HIV-antibody testing performance. Proceedings of 1995 Institute on Critical Issues in Health Laboratory Practice. Pioneers in Laboratory Research. Atlanta, Ga. Centers for Disease Control and Prevention 1996; Abstract.

Scornik JC et al. Multicenter Evaluation of tje Flow Cytometry T-Cell Crossmatch. Transplantation 1997;63:1440-1445.

Shehangian S. Proficiency Testing in Laboratory Medicine: Uses and Limitations. Arch Pathol Lab Med 1998;122:15-30.

Skendzel LP, Hanson DJ, Cirin W. The 1949 College of American Pathologists survey revisited. Am J Clin Pathol 1970;54:493-495.

Smith ML, Bronner WE, Shimomura ET, Levine BS, Froede RC. Quality Assurance in Drug Testing Laboratories. Clinics in Laboratory Medicine 1990;10:503-516.
Steindel SJ, Howanitz PJ, Renner SW. Reasons for Proficiency Testing Failures in Clinical Chemistry and Blood Gas Analysis. Arch Pathol Lab Med 1996;120:1094-1101.

Steadman GW. Survey of DNA Crime Laboratories, 1998. U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics 2000: 1-8.

Strand CL. Proficiency Testing: One Important Component of Continuous Quality Improvement. Am LClin Pathol 1994;90:393-394.

Stull TM, Hearn TL, Hancock JH, Handsfield JH, Collins CI. Variation in proficiency testing performance by testing site. JAMA 1998;279:463-467.

Sunderman FW. The History of Proficiency Testing/Quality Control. Clinical Chemistry 1992;38:1205-1209.

Technical Working Group on DNA Analysis Methods (TWGDAM). Guidelines for Quality Assurance Program for DNA Analysis. Crime Laboratory Digest 1995;22/2:21-43.

Tholen, Lawson NS, Cohen T, Gilmore B. Proficiency Test Performance and Experience With College of American Pathologists' Programs. Arch Pathol Lab Med 1995;119:307-311.

Thompson M, Wood R. International Harmonized Protocol for Proficiency Testing of (Chemical) Analytical Laboratories. J. AOAC Intl. 1993;76:926-935.

Thompson, WC. Subjective interpretation, laboratory error and the value of forensic DNA evidence: three case studies. Genetica 1995;96:153-168.

TWGDAM Quality Assurance Subcommittee and ASCLD-LAB DNA Proficiency Review Committee. Guidelines for DNA Proficiency Test Manufacturing and Reporting. Crime Laboratory Digest 1994;21:27-32.

US Congress. House and Senate Committees on the Judiciary. Subcommittee on Civil and Constitutional Rights and the Subcommittee on the Constitution. Forensic DNA Analysis. Hearings, June 13, 1991. No. 30. pp 83-315.

US Congress. House of Representatives. Committee on Small Business, Subcommittee on Regulation and Business Opportunities. Quality AIDS Testing. Hearings, Oct 19, 1987. Serial No. 100-32. 118pp.

US Congress. House of Representatives. Committee on Small Business, Subcommittee on Regulation and Business Opportunities. Deadly mistakes: are laboratory results reliable? Hearings, March 2 and May 3, 1988. No 100-43. 242pp.

US Congress. Senate Committee on Government Affairs. Subcommittee on Oversight of Government Management. Health Care financing Administration's management of medical laboratories. Hearings, March 23-24, 1988. No. 100-765, 796 pp.

US Congress. House of Representatives. Committee on Energy and Commerce, Subcommittee on Oversight and Investigations. Clinical laboratories. Hearings, March 10 and July 27, 1988. No. 100-146. 552 pp.

Wagner LR. The College of American Pathologists, 1946-1996: Laboratory Standards. Arch Pathol Lab Med 1997;121:536-541.

Welch MJ, Hertz HS. The How and Why of An Accuracy Base for Proficiency Testing. Arch Pathol Lab Med 1988;112:343-345.

Werret DJ. Remarks at Banbury Conference in DNA Technology and Forensic Science, Cold Spring Harbor, New York, November, 1988, p. 240.

Witte DL, VanNess SA, Angstadt DS, Pennell BJ. Errors, Mistakes, Blunders, Outliers, or Unacceptable Results: How Many? Clinical Chemistry 1997; 43: 1352-1356.

Yen-Lieberman B et al. Evaluation of a Quality Assurance Program for Quantitation of Human Immunodeficiency Virus Type 1 RNA in Plasma by the AIDS Clinical Trials Group Virology Laboratories. J Clin Micro 1996;34:2695-2701.

CASE CITATIONS

Daubert v. Merrell Dow Pharmaceticals, 113 S. Ct. 2786 (1993)

U.S. v. Porter, 1994 WL 742297 (D.C. Super. Ct., Nov 17, 1994).

Quality Assurance

New York v. Castro, Criminal Proceedings of the New York State Supreme Court, 144 Misc.2d 956, 545 N.Y.S. 2d 985 (Sup. Ct. 1989).

State v. Schwartz, 447 N.W. 2d 422 (Minn. 1989)

Statistics

Commonwealth v. Lanigan 413 Mass. 154, 596 N.E. 2d 311 (1992)

N.Y. v. Wesley 533 N.Y. 2d 643 (County Ct., Albany City, 1988)

State v. Anderson 853 P. 2d 135 (N.M. 1993), 115 N.M. 433.

State v. Mokit 579 N.Y.S. 2d 990 (1992)

State v. Vandebogart 136 N.H. 365, 616A.2d. 483 (N. H. 1992)

U.S. v. Porter 618 A.2d 629, D.C. App., 1992

U.S. v. Yee 174 F.R.D. 161 (N.D. Ohio 1991)

III: RESULTS OF SURVEYS OF FORENSIC DNA TESTING LABORATORIES, LAW ENFORCEMENT AGENCIES / CONDUIT LABORATORIES, DEFENSE ATTORNEYS, AND EXPERT WITNESSES

Two phases were conducted in this study. In Phase 1, laboratories performing forensic DNA testing in the U.S. were surveyed (see Appendices D1 & D2) and asked their willingness to participate in our study. Meanwhile, law enforcement agencies/conduit laboratories submitting DNA evidence to those laboratories were also surveyed (see Appendix D3). Laboratories and law enforcement agencies/conduit laboratories willing to participate in the study were asked to return the surveys and agreements (see Appendices E1, E2, & E3). 39 laboratories returned the agreements, while 63 law enforcement agencies and 9 conduit laboratories agreed to be submitting agencies.

In Phase 2, a new survey and the same "Agreement to Participate" form were distributed to the same laboratories as in Phase 1. Thirty laboratories that were willing to participate in the study returned the surveys and agreements. We also surveyed defense attorneys and expert witnesses, but due to the highly specialized and limited number of persons engaged in this work, we first obtained an initial list of defense attorneys and expert witnesses as recommended by members of our Review Panel, and then we employed a snowball sampling technique to expand the sample pool. In total, 6 defense attorneys and 11 expert witnesses returned this latter survey. We will examine the survey results in following sections of this chapter. In section A, the survey results of laboratories and law enforcement agencies/conduit laboratories in Phase 1 will be discussed. The survey results of laboratories and defense attorneys/expert witnesses in Phase 2 will be examined in section B and C respectively.

A. Results of Survey of Forensic DNA Testing Laboratories and Law Enforcement Agencies/Conduit Laboratories—Phase 1

1. Introduction

The primary goals of the survey of DNA testing laboratories (see Appendix D1) were to: a) identify all laboratories in the U.S. performing forensic DNA testing; b) gather descriptive information from these laboratories on procedures they follow in receiving and examining biological evidence, and reporting DNA test results; and c) determine the types of open and blind proficiency testing procedure in which they are engaged. The information we gathered on the receipt and processing of DNA evidence was necessary to devise candidate blind testing procedures. Detailed information on evidence collection, receipt, and processing was also required for us to consider the implications and requirements of introducing blind tests into forensic DNA labs on a national scale, and estimating corresponding costs. In addition, knowing the types of quality assurance procedures and proficiency tests DNA laboratories currently perform is an important piece of the total quality assurance/proficiency testing picture.

The survey was developed with various input from laboratory directors and analysts across the country, some of them members of our National Forensic DNA Review Panel and the DNA Advisory Board. The first draft of the survey was evaluated and pre-tested by fourteen laboratories across the country. These laboratories included state, local, and commercial laboratories.

Our goal was to survey every laboratory in the U.S. that was performing forensic DNA analysis. This listing was compiled initially from two sources: a National Institute of Justice (NIJ) DNA grantee list and the FBI's CODIS DNA Laboratory Survey conducted during summer 1995. A telephone call was made to every laboratory on the list to confirm that the laboratory was performing DNA analysis and to determine if there were any other laboratories in their state performing DNA analysis. From this list, a total of 151 laboratories were identified and surveyed in December 1996. One hundred two (102) surveys representing 42 states and 1 Federal agency were returned (68% return rate) and 94 laboratories indicated they were presently performing DNA testing on forensic case materials. Eight (8) laboratories indicated they would be performing such testing within the next 6-12 months. In contrast, the FBI's CODIS DNA laboratory surveyed 120 laboratories representing 42 states and 2 Federal agencies (See Table III-1).

With this survey, laboratories had the opportunity to indicate that they would be willing to participate in the actual feasibility segment of the study as potential test-target laboratories or as reference laboratories. Laboratories that were willing to do so were asked to return written agreements (see Appendix E1) that were then countersigned by the project directors and returned. The agreements set forth the terms of a laboratory's participation in the project. Fifty laboratories (49%) out of 102 respondents indicated on their survey they would be willing to participate in the study, and 38 (37%) actually returned signed agreements (We actually obtained a total 39 signed agreements because one laboratory returned a signed agreement but did not return a completed survey). We performed additional data analysis in which we compared the responses of laboratories agreeing to participate against those that indicated they did not wish to participate in this study. We found some interesting differences which will be further described in the narrative.

2. Demographics / General Characteristics of Respondents

Nine (9) laboratories responding to the survey indicated they send cases/evidence out for DNA testing. At the time of this survey, eight of these nine laboratories were not performing DNA testing. These laboratories indicated they either direct samples to: 1) an outside state or local forensic laboratory; 2) the FBI laboratory; or a 3) private or commercial laboratory with Roche and Cellmark the most commonly mentioned.

About 33% of the laboratories performing DNA testing are part of larger state forensic laboratory systems. Our respondents indicated that over three-quarters of these state systems also have other laboratories within the system performing DNA testing. We asked laboratories how many scientific personnel were performing DNA testing in their facilities. Over two-thirds of laboratories have 4 or fewer personnel engaged in DNA testing and while almost half the responding laboratories report 10 or fewer total scientific personnel, one-quarter report more than 25. Table III-2 also shows there is an average of 4.3 DNA analysts per laboratory and a mean of 19.4 total scientific personnel per laboratory for 93 laboratories responding to this question. Counties have fewer personnel while state facilities clearly have the most personnel. The reader should note that seven of the state laboratory responses included multiple laboratories throughout its state wide system.

Table III-1 Comparison of Survey Data

1995 FBI CODIS Survey	1996 UIC DNA Survey
120 labs surveyed	102 labs surveyed
42 States	42 States
2 Federal agencies	1 Federal agency
58 labs in 32 States performing RFLP	11 labs in 11 States performing RFLP
55 labs in 22 States performing PCR	42 labs in 24 States performing PCR
30 labs in 17 States performing RFLP+PCR	41 labs in 28 States performing RFLP+PCR
5 labs in 5 States performing STR analysis	27 labs in 15 States performing STR analysis

About forty percent (41) of the laboratories report they have been accredited by ASCLD (more than half of these accreditations have been granted since 1990, but only eight since 1994) and this is broken out by type of laboratory in Table III-3. We see a substantially larger percentage of state facilities has been accredited when compared with other types of laboratories. State laboratories and laboratories part of state systems responding to this survey are about twice as likely to be accredited as city or county facilities. Table III-3 also shows the percent of laboratories that have their DNA sections ASCLD accredited. We do, also, note that it is the laboratories with more analysts that have their DNA sections accredited. In addition, we found that state laboratories' DNA sections are more often accredited than other laboratory types. We inquired how many of their total complement of scientific personnel were ABC certified (See Table III-3). County laboratories tend to have fewer analysts certified than other laboratory types. Thirty-eight percent of laboratories indicated that at least one of their analysts was certified with an overall mean of 1.5 ABC certified scientists per laboratory.

Table III-2 - Average Number of Scientific/DNA Personnel by Laboratory Type*

Laboratory Type	Average Number of DNA Personnel	Average Total Number of Scientific Personnel
State** (n=45)	5.2	26.5
County (n=20)	2.7	15.6
City (n=15)	3.9	17.2
Independent/Private (n=11)	4.4	5.2
Other (n=2)	0	5.5
Overall Mean/Average	4.3	19.4

*The FBI laboratory is not include in this table, but reports 17 DNA analysts.
**Seven state laboratories, that are part of state-wide systems, answered for all laboratories within its system.

In comparing laboratories that were not willing to participate to laboratories that were willing to participate in our feasibility study with those that were, we found that those not willing to participate tended to have fewer scientific personnel performing DNA analysis, fewer ABC certified personnel, and a lower percentage of DNA sections ASCLD accredited (81% of those not consenting to participate versus 93% of those consenting to participate).

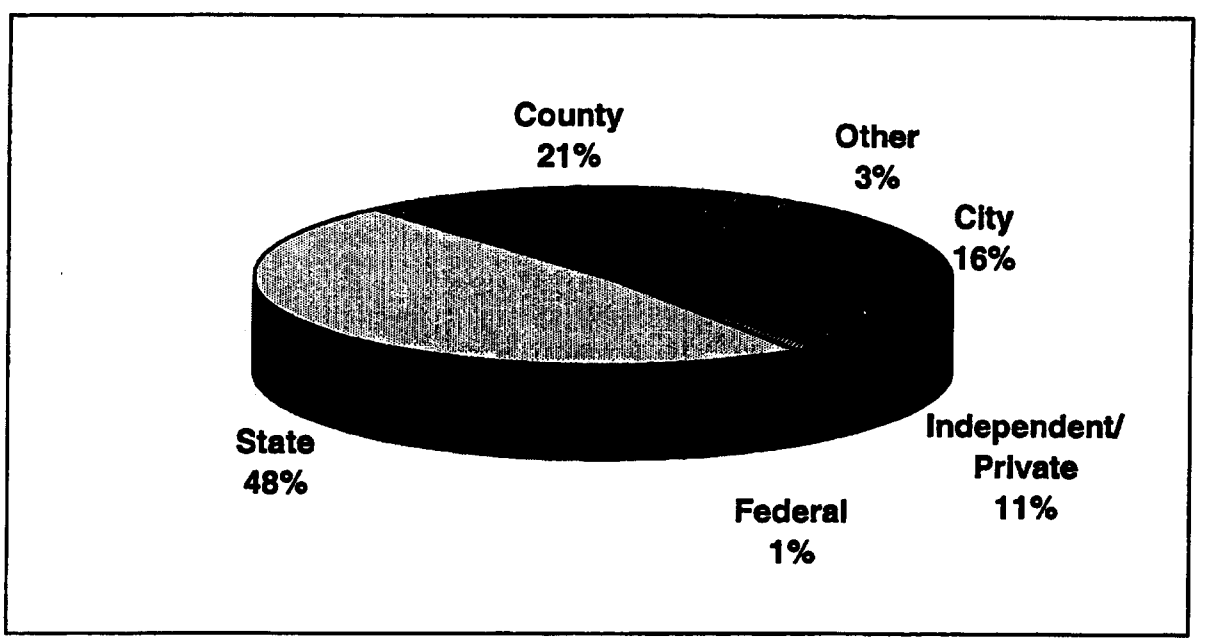
The number of scientists within laboratories needs to be taken into account in a blind proficiency testing program since in laboratories with several examiners there is little way to control which analyst receives a given test. If the objective is to test the proficiency of the laboratory, then which examiner performing the testing is not a major issue. If, however, the goal is to test specific members of the staff over a given time, many test specimens may have to be introduced to insure every examiner is tested. For example, it is conceivable multiple blind tests could be (randomly) assigned to the same analyst. Another compounding problem may be the division of effort in larger laboratories where more than one analyst is involved in the analysis of a single case specimen. Eleven of the laboratories responding to the questionnaire are private/commercial facilities, six are affiliated with a medical examiner's office, and another four are associated with a college or university. The majority (about two-thirds) of laboratories, however, are affiliated with law enforcement agencies.

Table III-3 - Accreditation of Facilities and Certification of Personnel by Laboratory Type

Laboratory Type	**Percent of Laboratories ASCLD Accredited**	**Percent of Laboratories' with ASCLD Accredited DNA Section**	**Average Number of Personnel ABC Certified**
State (n=45)	62	49	1.3
County (n=20)	30	30	1.9
City (n=15)	33	3	1.5
Independent/Private (n=11)	9	9	1.7
Other (n=2)	50	0	0
Overall Mean/Average	43	34	1.5

About 48% of the laboratories are units of state government, and 37% are associated with either city or county agencies (See Figure III-1). Interestingly, we determined that city and private/commercial laboratories were more willing to participate in this study than state and county laboratories. With the variety of laboratory affiliations and levels of government, each blind proficiency test would have to be designed and submitted in a manner consonant with the procedures and documentation used in the various jurisdictions. For example, commercial laboratories will analyze specimens on a fee per test basis and may only expect minimal communication, background information, and documentation. In contrast, a public/law enforcement laboratory may expect more extensive documentation and communication between analyst and police officer or district attorney. Our hypothesis to date is that the greater organizational and geographical distance between the submitting agency and the testing laboratory, the easier it will be to submit a blind test without it being detected.

Figure III-1 Types of Laboratories



We also asked laboratories the number of different agencies for which they provide DNA testing services. Of the 68 laboratories responding to this question, about 20% noted they serve only 1-2 agencies, and another 27% serve between 3 and 35 agencies. At the other end of the continuum 29% of laboratories report they serve more than 140 different agencies. We note that an additional 26 laboratories report they are unable to estimate the number of agencies they serve, and presumably service many agencies. In terms of developing blind proficiency tests for laboratories, it would be easier to test laboratories that serve more agencies. This allows the tester to select from a larger pool of agencies to submit blind PTs to the laboratory, thus permitting the use of a variety of case types from many different jurisdictions and conduit laboratories. Still, employing more agencies means the testing distributor must be familiar with

various protocols and submittal forms used, which could increase preparation time, effort, and costs accordingly. In contrast, with laboratories serving single agencies there is a greater likelihood the analyst would have acquaintances in the police agency, might become suspicious if they were not aware of the fictitious case described, and may have a greater tendency to make follow-up inquiries/telephone calls to the submitting officers. In terms of the populations served by these agency-clients, laboratories report about one-quarter serve populations less than .75 million, another quarter serve populations between .75 million and 1.2 million, and over half serve agencies with responsibility for areas exceeding 1.2 million population.

If the blind PT were to be a sexual assault case scenario, it would be important to determine what type of evidence collection kits are used and which specimens are absolutely required (i.e., whether failure to collect items called for by a kit would be automatic "flags"). Laboratories were queried if they had a standardized evidence collection kit for victims of sexual assaults, and 79% reported they did. About half of these kits are issued by the forensic laboratory itself, another 18% by the submitting police agency, and 12% by a victim services agency. Almost 60% of the evidence kits received by laboratories are custom made, either manufactured in-house by the laboratory or by an outside vendor. Another quarter of the kits are generic kits commercially produced. Forty-five percent of laboratories report they employ a standardized suspect evidence collection kit, of which 60% are furnished by the laboratory and 15% by the police agency. In sum, any organization engaged in nationwide blind proficiency testing would require access to such kits if they planned to distribute a sexual assault based scenario.

3. Evidence Collection Policies and Issues

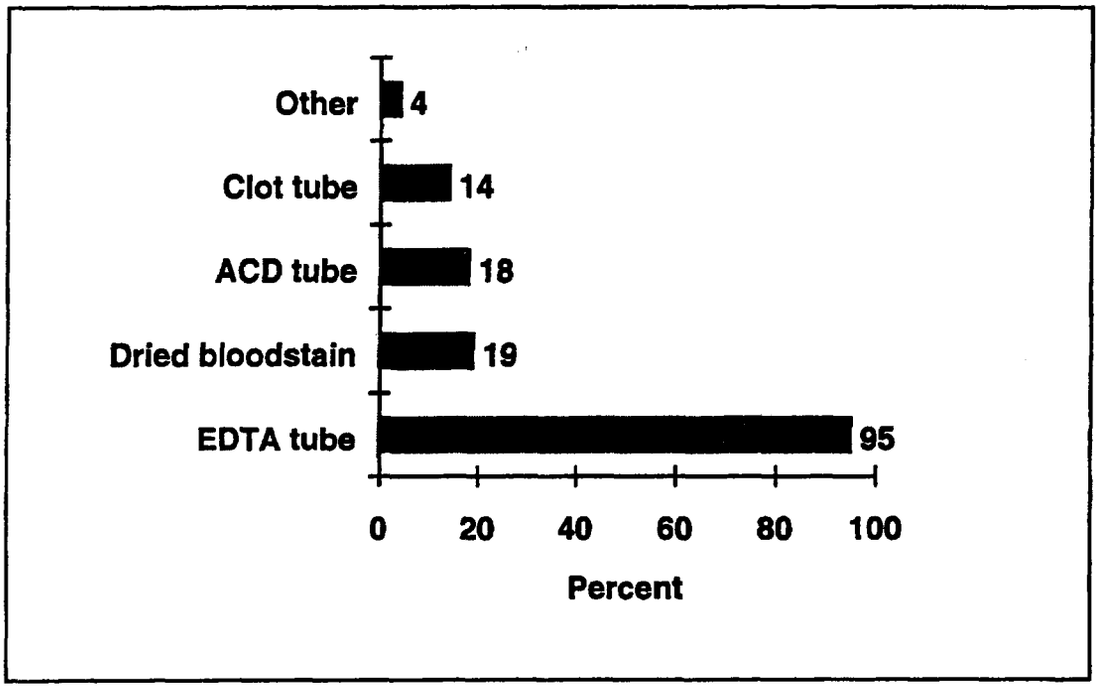
The manner in which specimens are preserved and received is important in the development and manufacture of a blind PT. Laboratories were asked in what form they received specimens from suspects in sexual assault cases (See Figure III-2). Most (95%) receive blood in EDTA (purple top) tubes. Between 15% and 20% of laboratories also accept blood either in ACD tubes, clot tubes, or in the form of bloodstains. In terms of saliva, the predominant collection media used are filter paper and swabs (between 40% and 50% use one, the other, or both). Less than 20% accept saliva samples on cotton, and less than 10% accept liquid saliva. In terms of head and pubic hair standards, about 95% use envelopes to hold the specimens.

The issue of specimen stability must be of concern to PT manufacturers if the laboratory requires a whole blood specimen or liquid saliva samples. Since liquid specimens may degrade, timing of specimen collection from donors must be closely correlated in time with case submission.

4. Biological Evidence Acceptance Policies

Sixty percent of laboratories reported that they accept <u>all</u> types of biological evidence. Of these, private/commercial laboratories more often accept and type all types of biological evidence than other laboratory types. Of the remaining 40% (39) of laboratories that do not accept all forms, one-quarter do not accept soft tissue or feces, about half do not accept bone, teeth or hair, and three-quarters do not accept urine.

Figure III-2. Form in Which Laboratory Receives Blood Specimens from Suspects in Sexual Assault Cases



About 80% of laboratories also report there are certain case circumstances under which they do not initiate the DNA typing of biological evidence. Table III-4 shows these responses and their frequencies. The leading situations are in cases of sexual assault where either known blood from the victim is not submitted or where vaginal swabs or semen are submitted without knowns from a suspect, and in all types of cases where questioned blood is submitted without known blood samples. About a third of laboratories also state they will not proceed with an analysis in various cases without suspects. Presumably this practice will change with the development of the CODIS databanking system which encourages laboratories to submit unknown suspect cases in hopes of linking this suspect to an offender in the database or another unsolved case. A blind PT submitting entity would have to be familiar with such decision policies for all the laboratories enrolled in the program.

5. Intake and Initial Processing of Evidence

In order to develop and successfully submit a blind PT to a laboratory, particular individuals within various submitting agencies (i.e., police departments, crime labs, medical examiner's office, and hospitals) need to be identified. Knowing the range of agencies and personnel who might be submitting DNA typeable evidence in different criminal case contexts is valuable in creating different scenarios and evidence types to challenge the DNA testing laboratory.

Table III-4 Circumstances Under Which Laboratories Will Not Proceed with DNA Analysis (n = 73 laboratories responding that there are such circumstances)

Circumstance	Percent
Missing "Knowns"	
In sexual assault cases if known from the victim was not submitted	63
In cases where questioned bloodstains are submitted without knowns	53
In sexual assault cases where vaginal swabs or semen are submitted without knowns from a suspected depositor	49
No Suspect(s)	
In a blood comparison case where there is no suspect	36
In a sexual assault case where there is no suspect	36
In any type of case where there is no suspect	34

The frequency with which various police and scientific personnel actually collect and package the DNA typeable material that is submitted to the laboratory for analysis is shown in Table III-5. Police evidence technicians, uniformed officers and detectives with special training, and regular detectives (in that order) were the most frequently cited personnel submitting such evidence. Crime lab personnel, civilian evidence technicians, and patrol officers were noted as the personnel least often involved in submitting such evidence.

Because evidence gathered from crime victims at hospitals is a common source of evidence submitted to DNA laboratories, we also asked which personnel most frequently collect biological evidence from victims. Nurses were the most commonly noted collectors of such evidence, followed closely by physicians. Physicians assistants and medical technicians were far less frequently involved in evidence gathering from victims. In terms of consenting suspects not under arrest, the most frequently cited collectors were nurses, followed by medical technicians and physicians. Physician's assistants, evidence technicians and crime laboratory personnel were only "occasionally to never" involved, with 6% of respondents saying these personnel were never involved in that type of situation. In terms of suspects actually under arrest, nurses were the most commonly noted collectors of biological evidence, followed next by medical technicians and physicians with physician's assistants, evidence technicians and crime laboratory personnel far less involved.

Table III-5 - Frequency Different Personnel Submit DNA Evidence to Laboratories (n=73)

Personnel Category	Mean Frequency
Police Evidence Collection Technicians	2.5
Uniformed Officers or Detectives with Special Training	2.5
Detectives	2.6
Crime Laboratory Personnel	2.9
Civilian Evidence Collection Technicians	3.0
Patrol Officer	3.2

Frequency: (1)Always, (2) Most of the time, (3) Occasionally, (4) Never

Once again, personnel who typically collect this evidence in particular jurisdictions must be noted so that the forms are completed properly. If a nurse's initials are typically found on evidence containers, then it will be necessary for the nurse to be included in the "manufacturing loop" and to have her initial the manufactured evidence. Manufacturers would need to secure these initials or signatures to insure authenticity and inclusion of this person in the manufacturing "loop" in the event the laboratory analyst makes contact with them.

Pathologists, themselves, most frequently collect the biological specimens for DNA testing from bodies in medical examiner and coroner's offices. Autopsy technicians and coroners/coroner's assistants are involved occasionally, with police and crime lab personnel almost never involved. If the blind PT manufacturing agency were to attempt to simulate homicide cases involving evidence recovered from a victim, consultation with a medical examiner or coroner personnel would be essential.

6. Receipt of Biological Evidence by the Forensic Laboratory

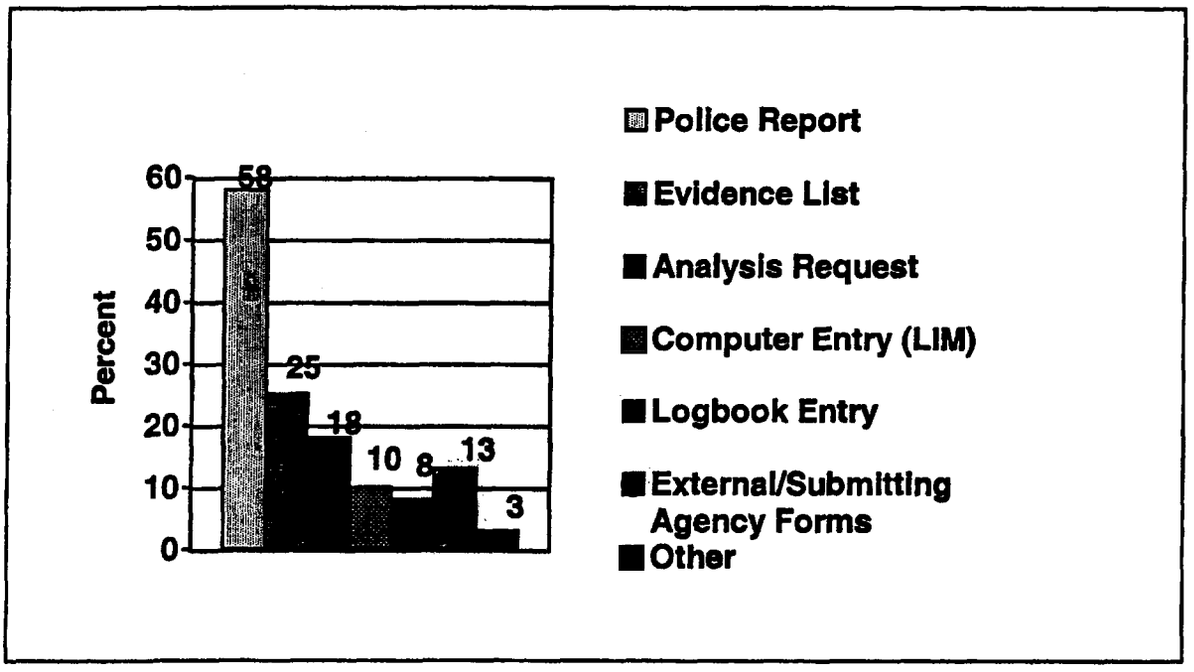
For blind PT introduction to be successful, the testing agency needs a thorough understanding of the laboratory's intake policies and paperwork. Any testing organization would need a detailed briefing by people familiar with the laboratory's practices before the testing process commenced.

The two most common types of personnel accepting biological evidence in forensic laboratories are evidence clerks (noted by 56% of laboratories responding to the questionnaire) and by a forensic examiner (42% of laboratories). Far less frequently (in less than 20% of laboratories) is such evidence picked up by a forensic examiner or clerk from a police evidence custodian. In about 94% of laboratories the biological evidence is received directly by a laboratory representative from a law enforcement agent. Practically two-thirds of laboratories also receive such evidence either through the mail or from UPS/FedEx or some other commercial carrier, and about 60% of respondents also report the forensic examiner may collect the biological evidence directly. Most often (three-quarters of laboratories), a receipt for the evidence is provided to the submitter in the form of a copy of the lab evidence submission form. Only about a quarter of laboratories supply a separate receipt document. Almost half (43%) of laboratories also require other types of documentation when the evidence is logged in. Of those, laboratories most often (58% of respondents) require a police case incident report (See Figure III-3). A minority of laboratories requires other miscellaneous reports; between 20% to 25% of laboratories also require an evidence/property list, a computer system (LIM), or some type of logbook entry. For those laboratories expecting a police report, a proficiency test distributor would need the cooperation of the relevant police agency to produce such a report.

Eighty-five percent of laboratories expect that various types of investigative information accompany the biological evidence upon submission, but there is little consensus among responding laboratories as to what type of information is required. A wide variety of reports/descriptions were noted, with the most frequently cited items being offense reports and related background information about the crime and about suspects and/or victims, but even these were noted only by about 10% to 15% of laboratories. Little information was expected in terms of particulars about sexual assaults. Only about 10% of respondents expected that requests for specific types of analysis accompany the evidence and 15% expected specific information concerning the relevance of evidence and knowns to the crime in question. Less than 10% of laboratories expect to have dialog with detectives in the case. The implications for a

nationwide testing program are that a detailed set of information report requirements must be constructed for every participating laboratory.

Figure III-3. Percent of Laboratories Requiring Different Documents with Evidence Submission



Laboratories that are part of larger state systems were asked the types of information / documents they expect from the submitting laboratory. Of the 28 laboratories responding to this question, the type of information most frequently requested (by 57% of the laboratories) are the serological results/reports of the initial testing lab. The next most commonly expected information was an evidence submission form (39%), followed by an analysis request form (29%). Laboratories are about evenly divided in terms of re-numbering/re-labeling items submitted and using the number/labels assigned by the submitter. Using a non-DNA-typing laboratory in a state system such as a "conduit" should be considered as one of the possible routes for submitting evidence for DNA blind proficiency testing to a target laboratory.

7. Prioritization and Case Assignment Policies

Given the rising caseload demands on forensic laboratories, we were interested in determining the extent of backlogs among laboratories and how they prioritized cases. A significant backlog could affect the timely return of blind PT results. Three-quarters (75%) of laboratories reported that there was at least a two week delay on average between the receipt of cases and the beginning of DNA typing. It is the larger laboratories with more total scientific staff and DNA analysts, that are more likely to have case backlogs. On the whole, county laboratories are less likely to have backlogs. Only a third, however, reported that they maintained a written policy that determined the priority given incoming cases. The absence of a written policy does not necessarily mean they don't have a policy. Similarly, less than 15%

reported they had a formal number or coding system for assigning case priorities. We determined that state laboratories are more likely to have prioritization policies and formally assign cases than other laboratory types.

Laboratories did report that the individuals within the laboratory assigning case priorities were generally the biology unit supervisor and DNA unit supervisor, which together constituted more than half of all replies. The third most commonly reported practice, noted by 17% of respondents, was for a combination of individuals to set priorities. Interviews with such supervisors would appear to be advisable in order to construct cases that have a reasonably good chance of getting worked. Investigators or other "users" should also be able to advise the types of cases that will be worked without delay.

It is very common (90% of laboratories citing this practice) for laboratories to assign cases a priority based on discussions with the police, and almost three-quarters in discussions with prosecutors (or defense counsel). In terms of blind proficiency testing, the organization issuing the blind tests might try to collaborate with these individuals to try to pressure the laboratory to process a blind PT more quickly. Such pressure, however, may also create suspicions the case is a "blind." Probing still further, laboratories were asked the importance of particular factors in giving cases/specimens priority. Clearly the most important factor cited was the date of trial, with almost 90% of laboratories citing this as "very" important. This was followed by the need for the results in securing an arrest warrant, the seriousness of the case, and to provide investigative leads—considered by over 90% of laboratories to be from "somewhat" to "very" important. To a lesser extent, a charging or preliminary hearing deadline, the date the evidence was received, and the willingness of a prosecutor to use the DNA results were noted as somewhat important priority factors. Interestingly, more than half the responding laboratories labeled the willingness of prosecutor to use DNA typing results if provided to be "unimportant." See Table III-6 which rank orders those factors considered very important.

8. Assignment of Cases / Specimens for Analysis

We also wanted to know how cases were assigned to particular DNA analysts. Although 39% of laboratories reported they had a "formal" system by which cases were assigned to analysts, an almost equal number (34%) noted a more informal system of rotation among analysts. Only 15% reported that assignments were random. Seventy-five percent of laboratories reported it was the DNA or serology unit supervisor who made such assignments. Less than a quarter reported assignments were made by a director or assistant director of the laboratory. In terms of internal blind proficiency testing, if a supervisor is responsible for case assignment, a blind PT can be targeted to a specific analyst within the laboratory. This would avoid the problem noted earlier about not reaching every examiner with a proficiency test, but would be compromising the other goal of preventing everyone in the laboratory from knowing a blind test was in progress.

In cases involving multiple biological evidence items, about three-quarters of laboratories state that they all are assigned to the same analyst. The same is true in cases where biological evidence items are subsampled: all such items are usually (75% of the time) directed to the same analyst.

Table III-6 Factors Considered Very Important in Setting Case Priority

Factor	Mean Frequency
Trial date (n= 79)	1.2
Needed for arrest warrant (n=77)	1.5
Seriousness of case (n=79)	1.6
Provide investigative leads (n=71)	1.7
Charging/preliminary hearing deadline (n=76)	1.8
Date evidence received (n=78)	2.1
Willingness of prosecutor to use DNA typing results if provided (n=69)	2.3

Importance: (1) Very important, (2) Somewhat important, (3) Unimportant

A follow up question was asked as to how it is decided which analysts receive different or subsamples of evidence items in a case. The most common method, noted by about 42% of respondents, was assignments are based on the availability of analysts, and secondarily (38%) on the specialization of particular analysts (PCR, RFLP, etc.). The implication for blind proficiency testing is that the case could be constructed such that it would be directed to a particular analyst or section.

9. Analysis of Evidence

Laboratories were asked if they performed conventional serological testing on biological evidence and 59% reported they did. Twenty-nine percent of laboratories reported they always performed such testing before DNA testing was performed and another 53% said it was sometimes done before DNA testing. Only 16% reported it was not performed. Laboratories reporting they performed conventional testing none or some of the time were asked to explain case circumstances in which they did. Three primary situations prevailed: 1) it depended on the quantity and quality of specimens; 2) the tests were used for screening purposes only; and 3) it was performed because of special requests and/or needs of the case. If one were to develop a blind PT under which these situations prevailed, the conventional serology test reference results would have to be ascertained to establish the complete array of values for comparison with the blind results.

Laboratories were surveyed for the number and different loci typed (both RFLP and PCR). This has helped to determine which loci are currently being used and thus what loci need to be typed by reference laboratories. Forty-four percent of respondents stated they performed both RFLP and PCR testing, 45% reported they performed PCR testing only, and 12% reported they performed only RFLP testing. This breakdown is different from the results reported in the 1995 CODIS survey in which only about 21% of laboratories were performing RFLP and PCR, 41% were performing RFLP, and 38% were performing PCR only. Consequently, we see that a much higher percent of laboratories (in the current survey) are currently performing both RFLP and PCR, a much lower percent are performing RFLP only, and a slightly higher percent are performing PCR only (See Table III-7).

It was also determined that city laboratories are less likely to do both RFLP and PCR, and more likely to do perform PCR only. Presumably, this is because city laboratories have less resources than larger state laboratories and therefore cannot engage in both RFLP and PCR analysis. In comparing laboratories that were not willing to participate to laboratories that were

willing to participate, we found that those not willing to participate tended to perform RFLP or PCR but not both.

A total of 51 laboratories reported performing RFLP tests. We asked a series of questions and broke out the responses according to laboratories performing only RFLP, and those performing RFLP and PCR. Initially, we combined these responses in Figure III-4 which displays the various loci that laboratories have the capability of typing and those they generally type. More than 90% of laboratories have the capability of performing D1S7, D2S44, D4S139, D5S110, and D10S28. About 75% of laboratories have the capability of typing D17S39, and only about 30% have the capability of typing D17S26 and D14S13. The corresponding percentage of laboratories that generally type these loci are also indicated in the Figure III-4.

Table III-7 Comparison of DNA Testing Capabilities (%)
1995 FBI CODIS Survey vs. 1996 UIC Blind PT Survey

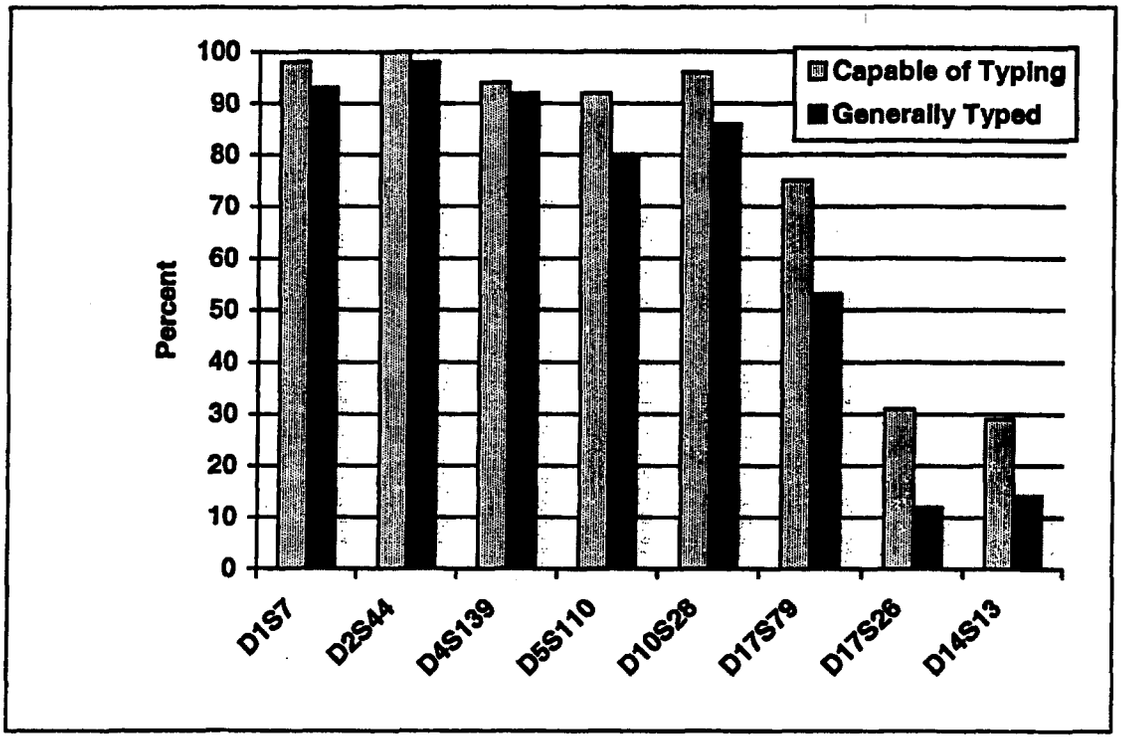
Capabilities	1995 CODIS Survey (n = 143)	1996 UIC Survey (n = 102)
RFLP only	41	12
PCR Only	38	45
RFLP and PCR	21	44

In terms of the 11 laboratories that perform only RFLP testing, they all reported they employ "FBI methods" or some modification of them. Two reported they have the capability of typing five loci, four reported six, four said seven, and one laboratory reported ten. In terms of how many loci they usually type, the most (half) reported five, and a quarter more indicated six. We also asked the particular loci they were capable of typing. All laboratories that do RFLP typing report that they have the capability of typing D1S7, D2S44, D4S139, D5S110, and D10S28, while three-quarters report that have the capability to type D17S79. All laboratories also report they generally type D1S7, D2S44, D4S139, and D10S28, and all but one generally type D5S110.

Five of the eleven laboratories performing only RFLP (46%) report they use chemiluminescence detection. All (100%) laboratories report that if the DNA is of insufficient quantity or quality for RFLP that they send the specimens elsewhere for PCR-based typing (About half of them send specimens to private/commercial laboratories for typing).

The next set of questions was directed to laboratories performing both RFLP and PCR-based DNA typing. When asked if they use PCR-based typing to screen evidence before deciding whether to do RFLP or additional PCR-based typing, only 7 of the 41 laboratories said yes, always, but another 49% said they did sometimes. HLA-DQA1 was the locus cited most frequently by those performing PCR for this type of screening, followed by PM and D1S80. Ninety-five percent of the 41 laboratories performing both RFLP and PCR-based typing follow FBI methods for RFLP. Fifty percent of laboratories report they have the capability of typing 6 loci, with 23% more stating they can type 7-8 loci, and 13% reporting they can type 9 or more.

Figure III-4. Percent of Laboratories with Different RFLP Typing Capabilities



The most frequently cited number of loci generally typed is 6, reported by 32% of the laboratories. Twenty-six percent state they generally type 5, and another 29% say they type between 1 and 4. Loci that 90% or more of laboratories report they have the capability of typing are D1S7, D2S44, D4S139, D10S28, and D5S110. D17S79 is done by more than 75% of this group. In terms of loci generally typed, 90% or more note D1S7, D2S44, and D4S139; in addition, D4S110 and D10S28 are generally typed by 80% or more of responding laboratories. Sixty-six percent of the laboratories performing both RFLP and PCR-based typing report they employ chemiluminescence detection.

For laboratories that perform PCR testing, Figure III-5 shows the loci laboratories have the capability of typing, both for PCR-only laboratories and for PCR-RFLP laboratories. Ninety-three percent have the capability of performing HLA-DQA1, 81% report that can perform AmpliType PM, and 58% can perform D1S80. Seven percent state they can perform amelogenin (XY) and 33% report they can perform STRs.

We also broke the data out for laboratories that perform only PCR testing and the results are similar. Ninety three percent (39) state they have HLA-DQA1 capacity, while 81% say they can perform AmpliType PM.

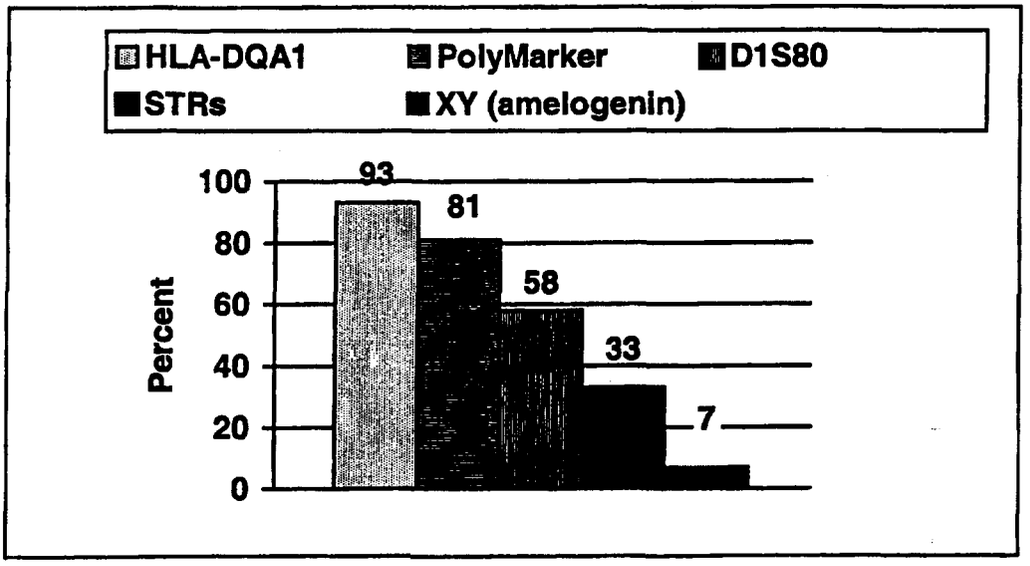
In laboratories performing PCR analyses, there were 27 laboratories that had the capability of performing short tandem repeat analysis. Figure III-6 displays these results and we see that the overwhelming majority (in excess of 90%) of laboratories do CSF1P0, TPOX, and THO1.

The three loci types next most often (in the 30% to 40% range) are vWF (41%), FESFPS (30%), and F13A01 (30%). For the twelve laboratories that perform PCR only and have STR testing capability, 100% of these laboratories report having the capability of typing the CSF1PO, TPOX, and THO1 loci. The next most frequently typed loci are vWF, FESFPS, and F13A01

performed by 42%, 25%, and 25% of the laboratories respectively. No other loci are typed by more than 20% of the laboratories performing STRs.

For laboratories that do both RFLP and PCR, 93% say they can perform HLA-DQA1, and 81% report AmpliType PM. Only 37% report they have STR typing capability. Of these laboratories, 90%-100% have the capability of typing THO1, CSF1PO and TPOX. An additional 30%-40% of the laboratories state they are capable of doing vWF, FESFPS, F13A01.

Figure III-5. Percent of Laboratories with Different PCR-Based Testing Capabilities



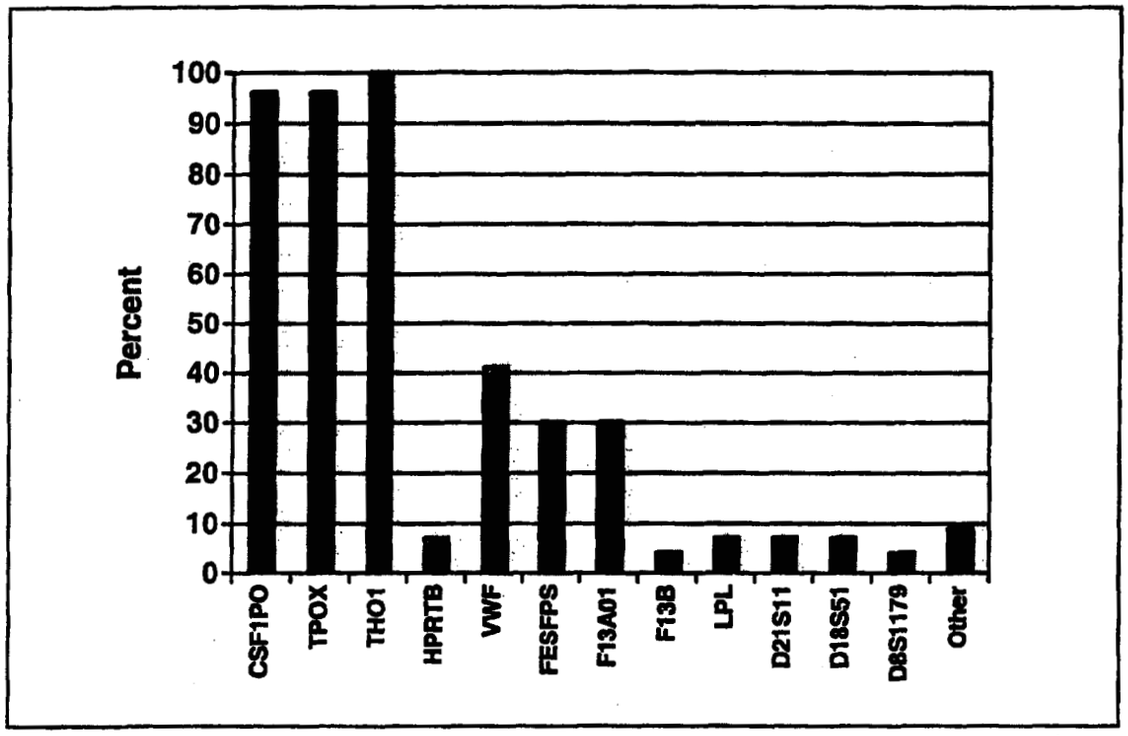
Blind proficiency testing suppliers/manufacturers must be aware of loci typed by laboratories to insure that reference laboratories type them as well for comparison.

10. Laboratory Notes and Repeat Testing

While external open proficiency tests are concerned principally with correct analytical responses, a blind proficiency test has the potential to examine the whole process by which cases and case specimens are processed including pre and post analytical procedures. Laboratory notes and reports can be examined from a QA/QC standpoint along with analytical results and may help to track the source of a problem or error when one occurs.

All laboratories report they keep lab notes for use in preparing written laboratory reports. More than 80% of the laboratories say these notes are a combination of free form and fill-in type forms. Ninety-eight percent of laboratories state these lab notes are reviewed by another analyst or supervisor.

Figure III-6 Percent of Laboratories with Capabilities for Typing Various STR Loci



A question that needs to be addressed in setting down the goals of any blind testing program is the desirability of reviewing the laboratory work notes as well as the formal report. If a goal of blind testing is to review the entire evidence processing system, then access to these documents would be necessary.

11. Re-Testing and Re-Analysis

Only 18% of laboratories report they re-test specimens to confirm results. Of those that do, 41% say the same examiner that performed the original typing also does the re-testing. Just 18% state that this testing is performed by a different examiner, and 41% state that the testing may be done by either the original examiner or a different examiner (See Figure III-7). We found that smaller laboratories, particularly those with 5 or fewer personnel are much more likely to re-test specimens than laboratories with more than 10 personnel.

Re-analysis of previously worked case specimens is considered by some to be a form of blind proficiency testing. Labs were asked if specimens that were worked are sometimes given to analysts to be re-tested as a QA/QC measure. Thirty-two percent replied that they did so. This is shown in Figure III-8, comparing it with other types of QA/QC measures including proficiency testing. Next they were asked the frequency with which case specimens that are tested are re-tested by another laboratory. Only 2% reported this was often done and 64% said this was done occasionally; 34% reported this was never done (See Figure III-9).

When asked if there had ever been a typing discrepancy between their results and those obtained by another laboratory on the same specimen, 8% reported there had been. About two-thirds of laboratories said this had never happened and another quarter said they didn't know.

When asked if they routinely save a sample of each specimen for possible future re-testing, 93% of labs reported they did. When queried what it was they saved, 88% replied they saved both the evidence specimen and extracted or amplified DNA from the specimen. Five percent reported they saved the specimen only, and seven percent said they saved only the extracted or amplified DNA (See Figure III-10).

Figure III-7. Which Analysts Are Involved in the Retesting of Specimens to Confirm Results
* 18% (n = 17) of Responding Laboratories Retest Specimens to Confirm results

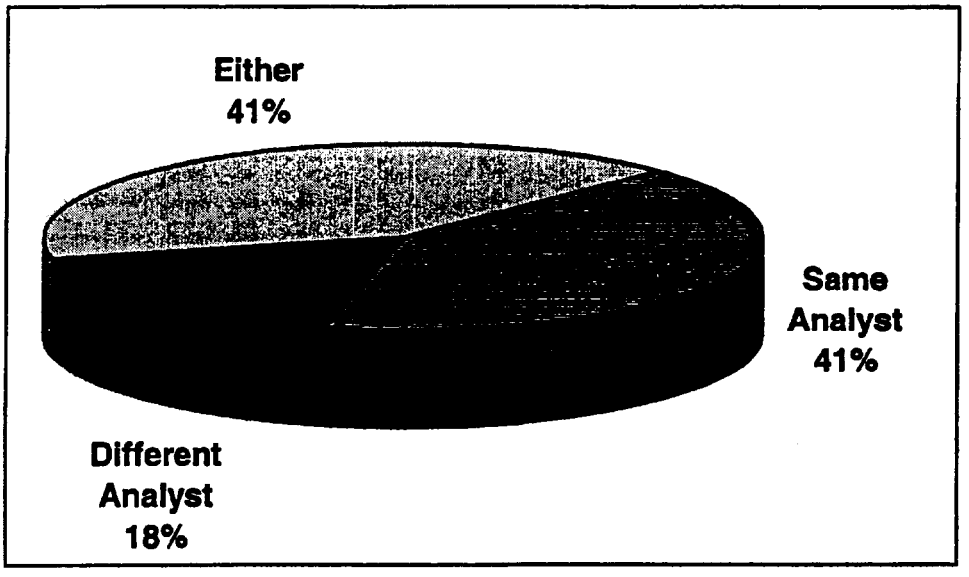
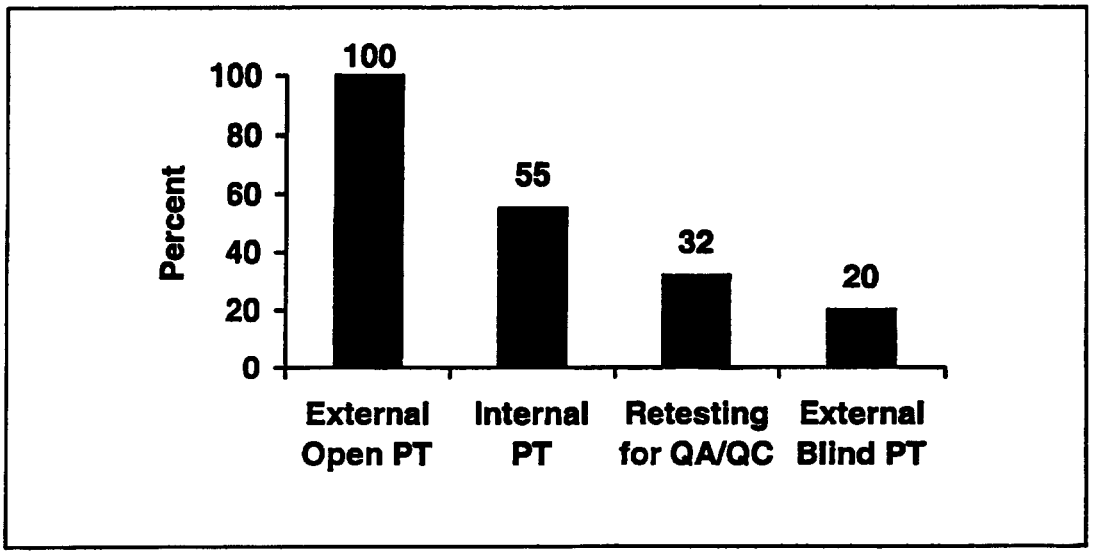


Figure III-8. Comparison of Different QA/QC Measures Utilized by Laboratories (n=94)

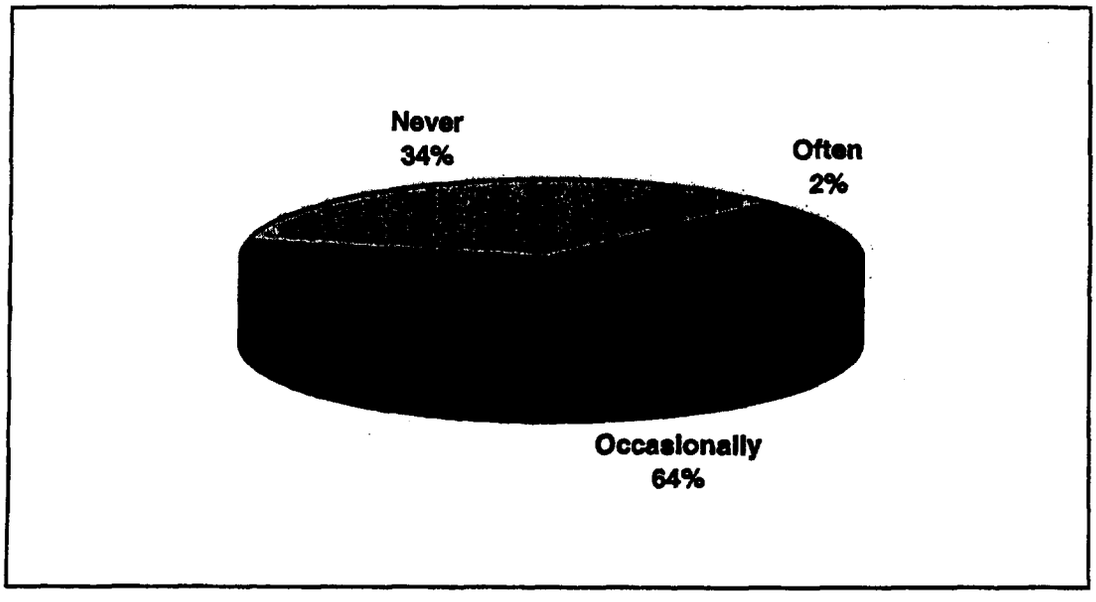


12. Laboratory Reports

We asked a series of questions concerning the content of laboratory reports issued. Only 10% of laboratories state they report RFLP band sizes for all specimens tested; 44% said they don't report band sizes at all. The remaining 46% of laboratories' RFLP reporting practices are somewhere in between. With respect to PCR testing, 83% said they report PCR locus types for all specimens, and only 4% said they do not report PCR locus types at all. The remaining thirteen percent report types for some specimens but not for others. Ninety-seven percent of the laboratories stated that their reports contain conclusions based on the typing results (in terms of which evidential specimens and knowns match or are of the same type). Ninety-two percent said they also include the frequencies of types found in the specimens broken out by ethnically distinct population groups in inclusionary cases. Reporting formats is an area that needs consideration in a national testing system to insure that results can be uniformly evaluated.

Ninety-seven percent of laboratories stated that reports are reviewed by another analyst or supervisor before they are issued. Only 80%, however, state that a supervisor must approve the report before it is sent to the requesting agency. More than 90% of laboratory reports are automatically sent to the submitting agency or the prosecutor (or defense, if a defense case) upon completion. When asked about informal reporting between the analysts and submitting agents, about 80% report that they frequently or always do so. Just 1% report they never give report results informally. Ninety-two percent state they give the results of their tests out over the telephone to the submitting agent or client, before the formal report is issued. For those laboratories that give results over the telephone, 45% of laboratories state they do this only when they are contacted. Another 38% say they do not make such calls routinely but will if it is an important case. This type of informal communication is a consideration in the operations of a blind proficiency testing program.

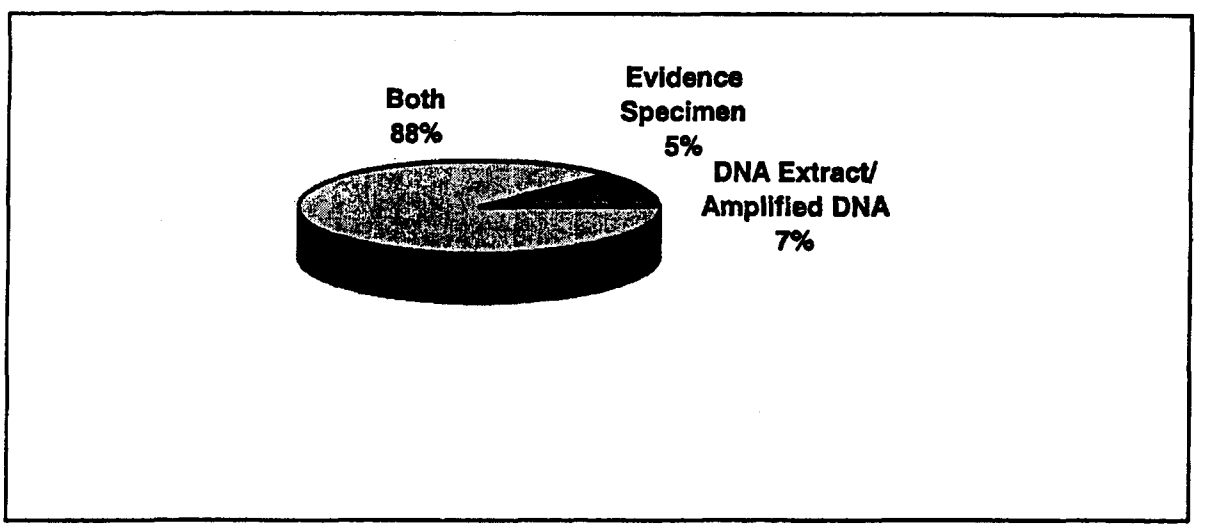
Figure III-9. Frequency Case Specimens are Retested by Another Laboratory



13. Databanking / CODIS

In many cases, the issuance of the laboratory report will not be the end of the blind PT. Many laboratories participate in DNA databanking or CODIS. For the most part, these labs are entering the DNA types of convicted sex offenders into the CODIS system. Many labs are also entering the DNA types of unknown suspects in sexual assault case into CODIS. The entry of this information into CODIS raises a series of other questions that need to be resolved before a national system of blind testing can be implemented. The fundamental point is that because these are fictitious cases it will be necessary to retract/delete such information once the laboratory is informed the case was actually a test.

Figure III-10. Items Saved for Possible Future Analysis



** 93% of laboratories responded that they routinely save specimens for future analysis (n=94)

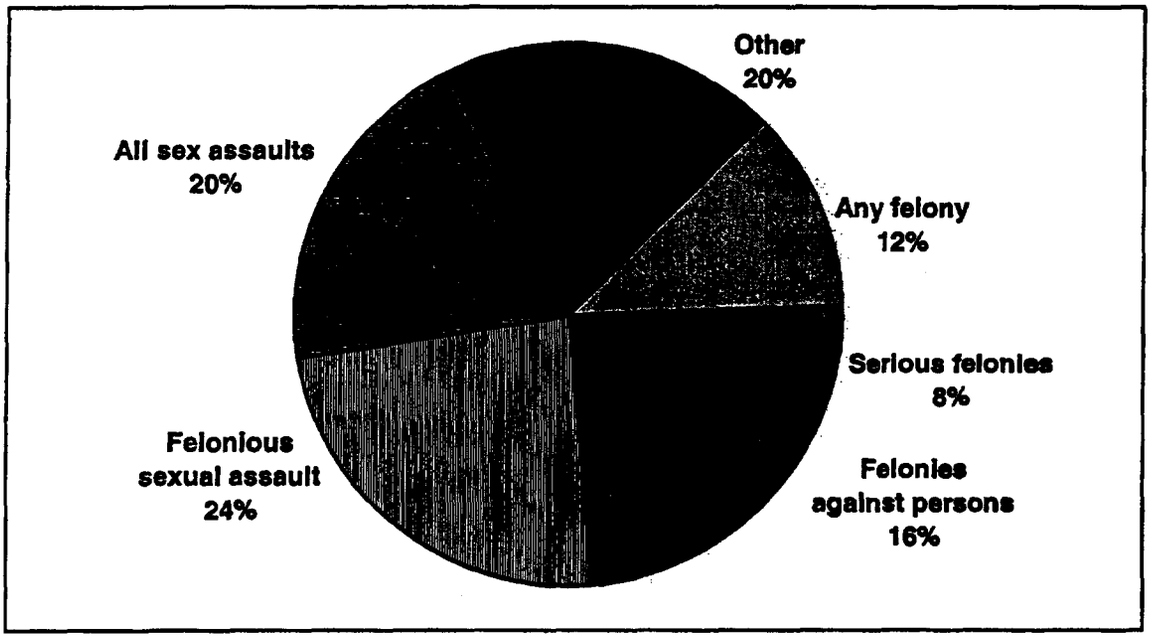
Twenty-eight percent of laboratories are involved in the processing of convicted offender specimens for entry into a statute-mandated databank. Of these laboratories, 80% are presently typing specimens. The databanking statutes governing this typing can cover a range of offense types depending on the state. See Figure III-11. Approximately fifty percent of laboratories state they are authorized to databank profiles for offenders convicted of some type of sexual offense. Sixteen percent say their statute covers any person convicted of a felony against persons, and another 8% state their statute covers any person convicted of a serious felony. Twelve percent report they type the DNA of persons convicted of <u>any</u> felony.

Ninety-two percent of laboratories engaged in databasing report that this database is maintained by their laboratory. The remaining 8% state the database is maintained by another laboratory either part of, or external to, their system. Eighty-nine percent enter the data directly into the database through a personal computer or terminal connection. The majority (60%) of laboratories report <u>any</u> analyst is authorized to enter such data. Another 16% state only a specified individual may enter data, while 8% more say data are <u>not</u> entered from their particular laboratory. Eighty-eight percent report that they can search the database from their facility using a PC. The balance of searching is performed from a central location that maintains the database.

In terms of personnel authorized to search the database, "any analyst" is the most common response (40%), followed by "specified individuals" (28%) and "supervisors" (16%).

Laboratories were also asked if the analysts who do casework also perform the databank specimen typing. Forty percent of respondents say the same analysts perform both functions, and forty-four percent report they have separate, dedicated CODIS analysts. The rest of the laboratories (16%) either rotate analysts or use another method. Although we are not addressing this issue in depth in this study, the blind testing of CODIS analysts presents an ancillary need in the forensic DNA typing community.

Figure III-11. Offense Types / Offenders Included in DNA Databank Statutes (n=25)



Blind testing of CODIS analysts is a more straight forward procedure which may be accomplished by making contact with the relevant criminal justice/correctional agency and simply have them introduce manufactured proficiency specimens. The unintentional submission of specimens from the same offender at different times by corrections officials can also serve as a "blind test" for CODIS analysts if the duplicate specimen gets typed and the duplication is detected by finding that the profile is already in the databank.

Fifteen percent of CODIS analysts are ABC certified. Of laboratories that have ABC certified CODIS analysts, one-third have 1-2, another third have 3, and the remaining third have 4 or more ABC certified analysts.

We inquired as to now many RFLP loci are databanked and the most common response (43% of laboratories) was 4 loci. An additional 14% reported 5, and 14% reported 6. Twenty-nine percent of the respondents indicated 3 or fewer loci are databanked. In terms of which RFLP loci are databanked, 100% reported D2S44, D4S139, and more than 80% database D5S110 and D10S28. For PCR-based loci, more than 50% database D1S80, and STRs. For the 5 laboratories typing STRs, 100% say they databank CSF1PO, TPOX, and THO1.

Eighty-five percent of laboratories databanking convicted-offender specimens collect whole blood in EDTA tubes. Sixty-nine percent of laboratories report that corrections personnel actually collect the specimens. Half state they provide written receipts when given these specimens, with more than 80% stating there are special forms that are used when collecting such CODIS specimens. More than 70% of all CODIS laboratories state that analysts are assigned such specimens randomly/on a rotating basis to keep a balanced workload. Approximately eighty-eight percent of CODIS respondents state they have a backlog of CODIS specimens awaiting analysis. For these laboratories with backlogs, half report that they analyze samples on a "first in - first out" basis. Another 26% say they prioritize samples based on the release dates of the offenders to be typed.

Eighty-four percent of CODIS laboratories say their results are reviewed by a supervisor or other analyst before they are entered into the database. In terms of repeat typing of CODIS specimens as a QA/QC measure, more than forty-five percent state that randomly selected specimens are checked; however 50% say there is no random, repeat testing. For those laboratories performing repeat typing, 25% say it is performed by a different examiner and 75% report it is accomplished by either the same or a different examiner. Three laboratories state they have no profiles in their database at the present time. Another third of the laboratories state they have between 20 and 2700 profiles. One-third more state they have between 3000 and 8700, and the balance (two labs), 1,100 to 12,000. Table III-8 compares the total number of databased samples in the 1995 FBI CODIS laboratory survey and the 1996 UIC Blind PT DNA survey. In a one year period, it shows an increase from approximately 50,000 samples to 75,000 total samples.

Laboratories involved in databanking and databasing enter the DNA types of unknown subjects from sexual assaults into the CODIS system. In terms of blind PT, it must be assumed that sexual assault blind PT cases will result in the suspect's profiles being databased. If blind testing is attempted on a large scale, CODIS system entry might create problems. Recruiting a sufficient number of suitable donors to prevent multiple matches in the CODIS system might be necessary. a succession of "hits" would probably soon lead the laboratories to decide the case from which the specimen came was probably a proficiency test and could cause other laboratories to be alerted to watch for blind tests in their as yet unworked cases. Identical specimens in blind PTs in several CODIS laboratories would probably not create any problem and could serve the additional purpose (have the additional value) of "testing" the CODIS system. Should a point be reached where interjurisdictional case to case cold CODIS hits are relatively common, fewer donors would be required for the construction of parallel blinds in multiple laboratories. Another issue is the need to purge the blind PT donor DNA profile from the CODIS system. Although the proficiency testing organization may attempt to see that all donors' DNA types are purged there cannot be an absolute guarantee to the donor that this will occur.

Table III-8. Number of Databased Samples in Nationwide Tally of CODIS Labs

1995 FBI CODIS Survey	1996 UIC DNA Survey
• 25 labs reporting • 46,244 RFLP samples • 3,000 PCR samples • 1300 RFLP+PCR samples • Total of **50,544** databased samples	• 21 labs reporting • Total of **76,977** databased samples (RFLP + PCR)

14. QA/QC Programs and Proficiency Testing

Seventy-one percent of the laboratories report they follow TWGDAM QA/QC guidelines, 6% state they have their own QA/QC programs, and 22% report they employ both. Figure III-8 compares the different types of proficiency testing done by laboratories. In terms of internal proficiency testing, 55% of laboratories report casework examiners perform proficiency tests and 45% do not. For those that do, 30% say that examiners perform one per year, 31% say they do 2-3 per year, and 15% say 4 or more. Another 16% state the number performed varies, and 11% report internal proficiency tests are administered only during training. Table III-9 displays an overall mean of 3.5 internal PTs/examiner/year.

Table III-9. Average Number of Proficiency Tests / Examiner / Year by Lab Type

Laboratory Type	Internal PTs	External Open PTs	External Blind PTs
State	3.3	2.2	1.8
County	5.4	2.6	1
City	1	2.3	2
Independent / Private	2	2.3	1.5
Other	4	3.5	3
Overall Mean / Avg	**3.5**	**2.3**	**1.7**

*Note: The "n" of laboratory types vary in each category of PT.

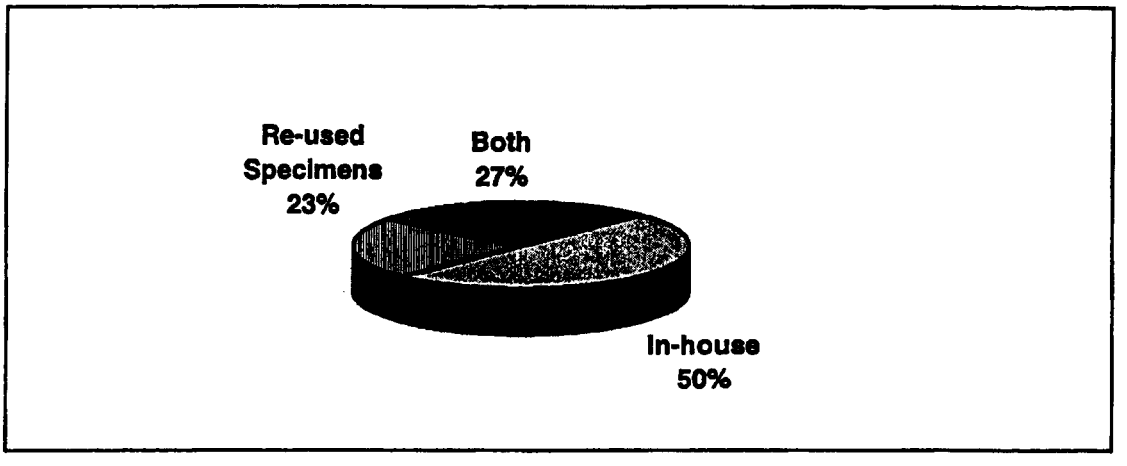
We also found that state laboratories are more likely to perform internal PTs while city laboratories are less likely to do them. In addition, laboratories that were not willing to participate in the study tended to perform less internal proficiency testing (46% of laboratories not willing to participate performed internal PTs compared to 63% of laboratories agreeing to participate).

Fifty percent of laboratories engaged in internal proficiency testing state they manufacture the tests in-house. Another 23% say they either use specimens from a previously analyzed case or commercially available external PT kit, and the balance (27%) state they do both (See Figure III-12). Some of these tests are not considered blind because the analyst presumably knows the specimens are from either a previously analyzed case or external PT.

What has not been discussed to this point is blind internal proficiency testing where the test is manufactured in-house. This model involves at least one person in the laboratory (e.g., QA/QC coordinator, supervisor) setting up a blind PT and introducing it as casework.

When the PT results are reviewed with the examiner by a supervisor or QA/QC coordinator, about half report that only the DNA testing results are discussed, and the other half state the test results plus evidence receipt, handling, and reporting of results are also discussed. In terms of dedicated CODIS analysts, half the laboratories report they participate in the same PT program as casework scientists, while the other half has a separate internal PT program for CODIS analysts.

Figure III-12. Type of Internal Proficiency Tests (n=52)

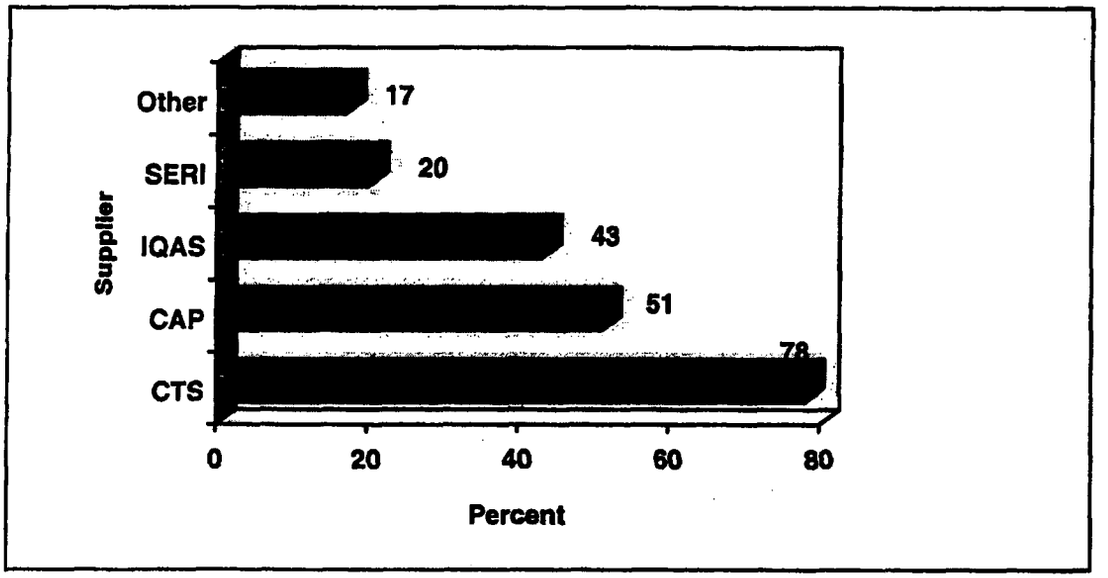


** 55% of laboratories reported performing some type of internal proficiency testing (n=94)

Every laboratory states that its casework examiners perform external, open proficiency tests. The great majority (70%) of laboratories report examiners perform two proficiency tests per year. Table III-9 displays the average number of external open PTs/examiner/year with an overall mean of 2.3. Figure III-13 shows the different providers of external open PTs. Collaborative Testing Services (CTS) is the most commonly mentioned provider (78% of labs). The next most frequently mentioned supplier (51% of labs) is the College of American Pathologists (CAP). Cellmark Diagnostics (IQAS) and Serological Research Institute (SERI) are indicated by 43% and 17% of laboratories respectively. If a nationwide blind proficiency testing program were introduced, these suppliers would presumably be candidates for manufacturing specimens and/or cases for it. In terms of dedicated CODIS analysts, about three-quarters of laboratories participating in CODIS and responding to this question, do not have a separate external proficiency program for these analysts, and one-quarter do.

Just 20% of laboratories report their casework examiners perform external blind proficiency tests with state laboratories more likely to engage in them. Of the 19 laboratories that do, only 9 (47%) state that these tests are issued "regularly". For those laboratories that say they perform these tests regularly, most do one test per year. The "occasional" users of blind proficiency tests also do them about once per year. The overall mean of external blind PTs/examiner/year for the 19 laboratories that perform this is 1.7 as shown in Table III-9. External, blind proficiency tests are most often made up in-house or by another forensic laboratory. None of the laboratories has a separate, external blind proficiency test program for CODIS analysts.

Figure III-13. Providers of External Proficiency Tests (n=94)



We also found the laboratories that were not willing to participate in the study tended to perform less external blind proficiency testing (14% of laboratories not willing to participate performed external blind PTs compared to 27% of laboratories agreeing to participate).

When laboratories were asked about their experience with external blind proficiency tests, most reported them successful. Eighty percent of those laboratories with experience in proficiency tests indicated supervisors informed analysts when they had been proficiency tested and what the results of the testing were. When probed further about blind proficiency testing, almost one-half replied either that blind PT was beneficial but difficult to implement, and/or that various logistical issues concerned them. When asked about these logistical/implementation problems, the issue cited most often (by a third of those responding) was their concern over communication and the potential breach of trust between lab analysts and law enforcement agents. The next two most common issues raised were the difficulties in deceiving experienced analysts, and issues surrounding how to flag, purge, or otherwise handle documentation (reports, notes, computer records) that resulted from blind proficiency tests.

We then asked laboratories if they were to participate in a fully blind DNA proficiency test and the fictitious case went through the laboratory undetected, what types of paper documents and/or electronic files would be created and thus potentially have to be purged. This was an open-ended question and only a minority of laboratories cited the types of documents (laboratory reports, notes, case files, evidence submission forms) that might have to be purged. Those types of items referenced most often were electronic files, with 27 laboratories specifically mentioning either CODIS or other DNA databasing information. a total of 20 laboratories mentioned they thought purging of such information would be a problem, but 30 who did not believe the purging would present a problem.

15. Survey of Law Enforcement Agencies / Conduit Laboratories

An integral part of blind proficiency testing procedures is the successful (and undetected) submission of actual proficiency test materials to the target laboratory through a law enforcement agency. The primary goal of the survey of law enforcement agencies and conduit laboratories was to determine which law enforcement agencies and/or conduit laboratories would be willing to submit evidence to a laboratory for DNA analysis. Table III-10 summarizes the data generated from these surveys. a total of 39 DNA testing laboratories that were willing to participate in the feasibility study were identified. These laboratories were surveyed and each asked to identify ten law enforcement agencies and/or conduit laboratories that submitted evidence to them for DNA analysis. One hundred sixty seven law enforcement agencies and twenty-two conduit laboratories were identified. Subsequently, these agencies and laboratories were surveyed and asked if they would participate in this study as a submitting agency. Sixty-three law enforcement agencies and nine laboratories agreed to be submitting agencies.

Law enforcement agencies that declined participation cited the following reasons: (1) the mock case paperwork that would be required would constitute a violation of departmental procedures or statutory prohibitions against filing knowingly false official statements; (2) insufficient personnel and time, and (3) would require explicit permission from the laboratory before making a decision.

Conduit laboratories were also queried as to the types of cases and/or situations in which they would submit evidence to another laboratory for DNA testing. The main reasons cited were: (1) laboratory does not perform DNA analysis; (2) for supplemental analysis of cases; (3) for paternity cases; (4) when laboratory resources are stretched, and (5) for re-analysis of casework samples for QC purposes.

16. Conclusions from Phase 1 Surveys

The survey results of the DNA testing laboratories and law enforcement agencies / conduit laboratories were important in the development of our candidate blind proficiency testing procedures. In addition, the data generated from the survey has given us a detailed picture of the characteristics of DNA laboratories across the nation, and most importantly, helped us to determine the types of quality assurance and proficiency testing procedures utilized in DNA laboratories.

Table III-10. Law Enforcement Agency / Conduit Laboratory Survey (June 1997)

- 102 laboratories surveyed (December 1996)
- 39 laboratories consented to be participants (surveyed May 1997)
- These labs identified 189 law enforcement agencies and/or conduit labs who submit evidence for DNA analysis (surveyed June 1997)
- 88 surveys returned
- 75 surveyes returned from law enforcement agencies
- 63 law enforcement agencies agreed to be potential submitters (84%)
- 13 surveyes returned from conduit laboratories
- 9 conduit laboratories agreed to be potential submitters (69%)

Although a great deal of useful data were generated from the survey, we also discovered that an equally important resource are individuals within submitting agencies (e.g., police officers). Police personnel who submit evidence for DNA typing are an integral part of any blind proficiency testing procedure. Not only do these officers submit the blind PT to the laboratory, they also supply insight as to the types of cases that would (or would not) be worked by a particular laboratory in a timely fashion, and the best way to construct cases so as not to arouse suspicion.

B. Results of Survey of Forensic DNA Laboratories—Phase 2

In December 1998, 137 Phase 2 DNA Laboratory Survey Instruments were mailed out to forensic DNA-testing laboratories across the nation. The principal purpose of this survey was to: (1) gather additional information from DNA-testing laboratories about their sample retention practices; (2) examine the types of internal reviews performed on completed cases, including re-analysis of evidence; and (3) evaluate the range of external reviews performed on completed cases by auditing groups and defense experts.

In addition, we inquired about defense discovery practices, i.e. the extent to which laboratory data were disclosed to defendants and subsequently reviewed (and possibly re-tested) by defense DNA experts. Out of the 91 labs that responded to our survey (66% response rate), 67 (74% of respondents) indicated that they are currently performing DNA analyses. The following results are based upon data returned by these 67 laboratories.

1. General Lab Information

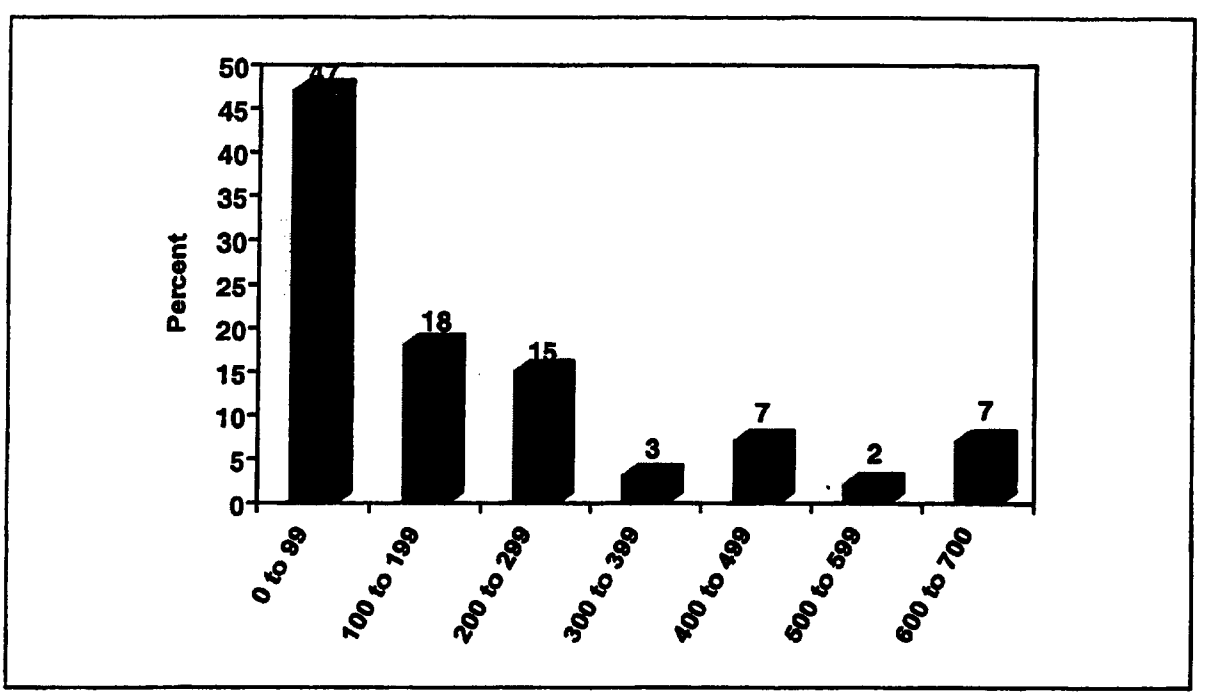
In terms of the geographical distribution of the 67 responding laboratories, 60% of the labs identified themselves as organizationally located at the state level. Twenty-four percent are at the county level, 12% at the city level, and 4% (n=3) indicated other organizational locations. No federal laboratories responded. Of all the surveyed labs, 61% indicated that their laboratories are ASCLD/LAB accredited, while 39% have not been accredited as of 1997. In a finer breakdown, 65% of the responding state labs are accredited, 56% of the county labs are accredited, 63% of the city labs are accredited, and 33% (n=1) of the remaining labs are accredited.

Furthermore, laboratories were questioned regarding the types of testing they perform, specifically: RFLP, Dot-Blots and/or D1S80, and STRs. Of the responding laboratories, 46% of the labs performed RFLP, 75% perform Dot-Blots/C1S80, and 69% perform STRs. (Percentages overlap because laboratories perform multiple types of testing.) Only one responding lab cited a testing technique other than the above mentioned three.

Twenty percent of the labs surveyed indicated that they perform all three types of testing. Fifteen percent of surveyed labs perform both RFLP and Dot-Blots/D1S80, 27% perform both Dot-Blots/D1S80 and STRs, and 9% perform both RFLP and STRs. Two out of the 67 responding labs (3%) perform RFLP only, 13% perform Dot-Blots/D1S80 only, and 13% perform STRs only.

In calendar year 1997, the responding labs analyzed and reported a total of 10,925 DNA cases. The number of reported DNA cases range from 6 to 700 cases per lab. The average value was 176 cases per laboratory, and the median was 108 cases. Approximately half of the responding labs had analyzed and reported fewer than 100 DNA cases (See Figure III-14).

Figure III-14. 1997 Analyzed DNA Cases (n=62)

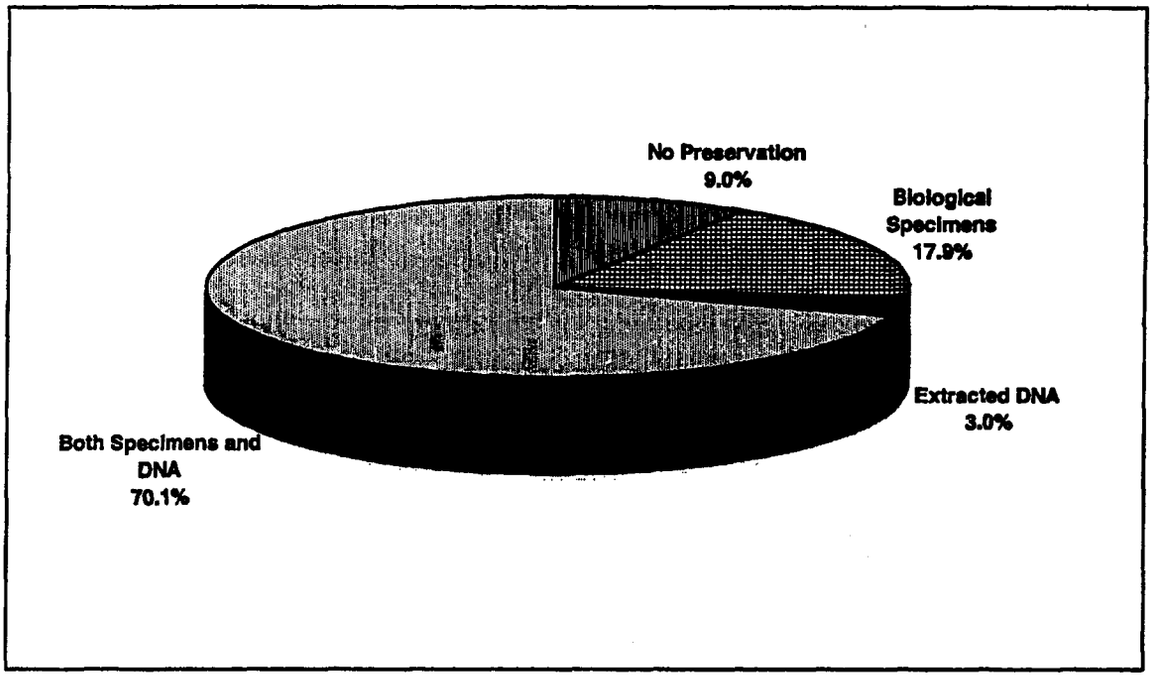


2. Sample / Evidence Retention

Prior to examining the internal and external re-testing procedures of the surveyed labs, laboratories were first asked about their retention practices, since they play a crucial step in determining the possibility of retests. First, laboratories were asked about the preservation of specimens for subsequent confirmation or retesting. According to the surveyed laboratories, 70% indicated that they preserve portions of the biological specimen and extracted DNA, 18% of the labs responded that they attempt to preserve a portion of the biological specimen only, and 3%

(n=2) attempted to preserve only extracted DNA. Nine percent (n=6) of the labs indicated that they do not preserve the specimen as a matter of practice (See Figure III-15). Of the six labs that do not preserve specimens for subsequent confirmation or re-testing, all are state laboratories.

Figure III-15. Evidence Retention in Laboratories (n=67)



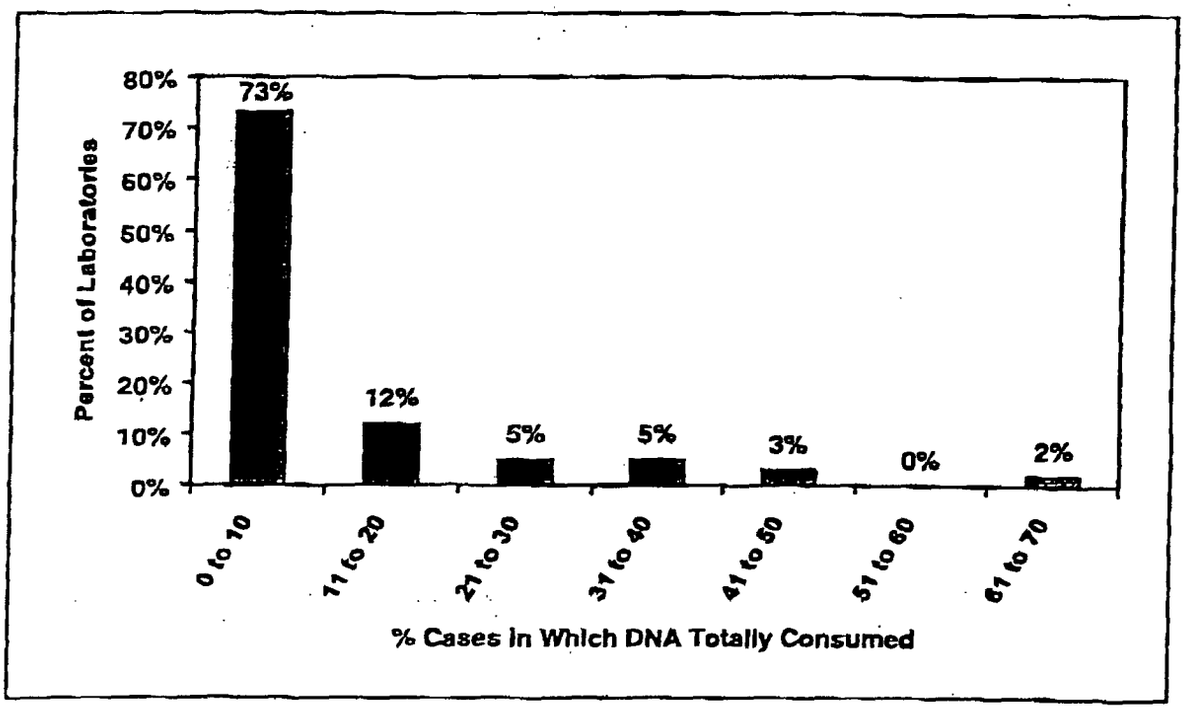
After DNA analysis on a case, 12% (n=8) of the laboratories indicated that they retain the original evidence (e.g., the original item or garment containing the biological stain). While four of these labs (50%) retain the evidence from between one-half year up to five years, the other four labs (the remaining 50%) retain the evidence indefinitely.

Although the rate of retention of the original evidence is low, 72% of all laboratories indicated that they retain a "cutting" of the evidence containing the stain of interest, while approximately 27% of the all responding laboratories retain neither the original evidence nor a cutting of the evidence. Of the labs that keep a "cutting" of the sample, 75% indicated keeping the cutting indefinitely, and the remaining 25% of the labs keep the cuttings from one-half year up to 10 years.

While retention rates of both accredited labs and non-accredited labs vary, accredited labs, in general, reported a higher retention rate than non-accredited labs. Seventy-eight percent of accredited labs reported to retain "cuttings" of the evidence containing the stain of interest, while only 61% of the non-accredited labs did. Besides sample retention, laboratories were also asked to estimate the percentage of DNA cases analyzed during 1997 in which key biological samples were so small that they were consumed in their entirety. Of the responding labs, approximately 73% reported that 10% or less of all DNA cases involved total consumption of key samples (with a median at 5%) while the remaining 27% of respondents reported that between 11% and 70% of

DNA cases analyzed in 1997 involved total consumption of key biological samples (See Figure III-16).

Figure III-16. Total Consumption of Biological Evidence (n=64)



After testing is completed, 82% of the respondents indicated that the evidence was returned to the submitting agency. Another 13% of labs indicated that the evidence was sent to the police evidence section/property division, while the remaining 5 % indicated other means of disposition. While majority of state and county labs returned their evidence to the submitting agency, majority of city labs (88%) reported to have sent the evidence to the police evidence section/property division.

Laboratories were then asked if they knew how long the evidence would remain in storage after the lab returns the evidence to the storage agency, and 52% reported that they did not know. Eight percent stated that the evidence is likely to be kept at the agency indefinitely, 14% believed that retention practices varied by agency and/or offense type, and 15% believed that the evidence would be kept at the agency until adjudication/court order. Furthermore, two labs (3%) indicated that the evidence is likely to be kept at their specified agency until the statute of limitations for the particular offense expires. The remaining 8% of the laboratories indicated no specific terms of retention.

In general, laboratories that returned evidence to their own police evidence section/property division had better retention knowledge than laboratories that returned the evidence to the submitting agency. While 60% of the labs that return the original evidence to the submitting agencies did not know how long it would be retained there, a much lower percent

(11%) of the labs that return the evidence to their police evidence section/property division had no knowledge of retention time. Additionally, while less than 5% of labs that return evidence to the submitting agency indicated that evidence is kept indefinitely, a higher percentage of labs that return the evidence to their own police evidence section/property division believed that the evidence would be held indefinitely.

Evidence retention is a key factor in assessing realistic alternatives to blind proficiency testing because it directly influences the ability to re-test evidence already analyzed. According to responding labs, the likelihood of a case being re-tested is not only based on the availability of the sample and the means used to select the sample (random or otherwise), but also whether or not there is enough retained sample to be re-tested. Our findings above indicated that approximately one-third of the laboratories do not retain the original evidence, nor a cutting of it. Furthermore, 25% of the labs also indicated that between 15 and 70% of their DNA cases resulted in total consumption of key biological samples, meaning those cases will not be re-tested.

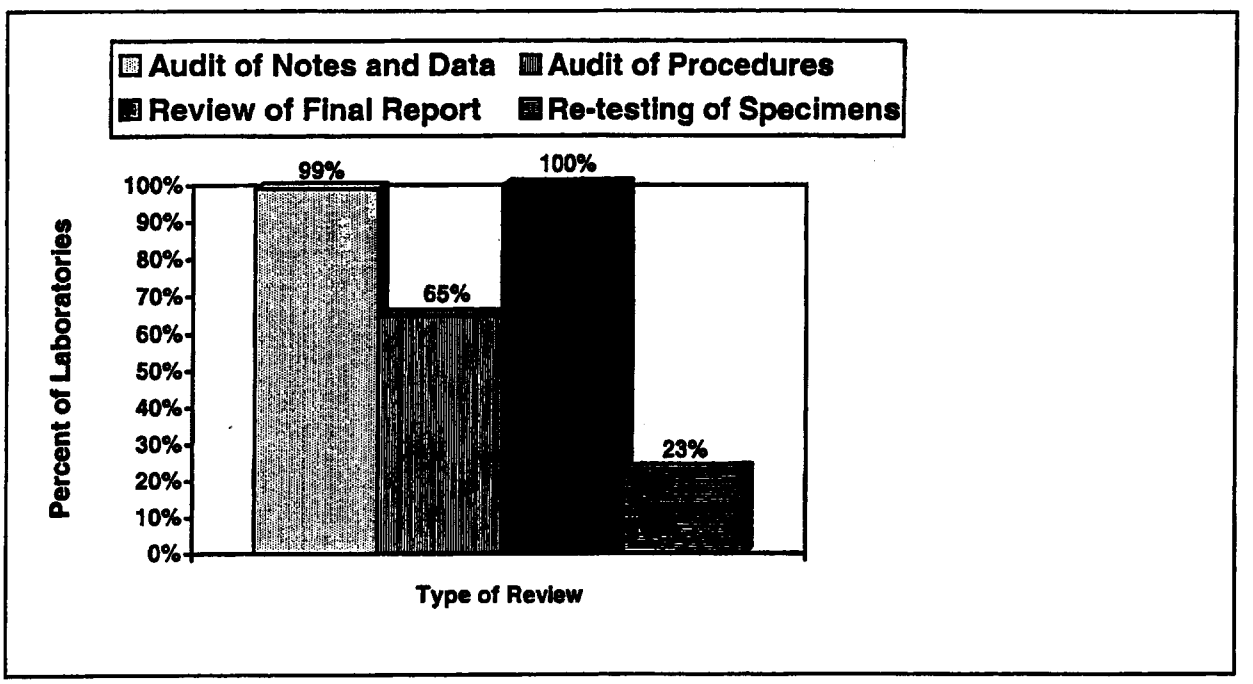
3. Internal Review / Reanalysis

Because we are aware of QA/QC standards and procedures employed by ASCLD, TWGDAM, and the DAB regarding technical reviews and audits, the survey instrument sought to explore the frequency, nature, and outcomes of these reviews. During calendar year 1997, almost all of the responding laboratories (99%) estimated that between 95 and 100 percent of their analyzed DNA cases were subjected to an additional internal review of laboratory results (for QA/QC purposes) by their laboratory peers, supervisors, quality assurance personnel, or other members of their parent agency. One laboratory reported no internal review on analyzed DNA cases.

Of the laboratories that conducted internal reviews on analyzed DNA cases, all labs indicated a review of the final report. Almost all of the labs (99%) indicated reviews of notes and data, 65% indicated auditing of procedures (approximately 2/3 of the laboratories that audited procedures indicated that such audit was performed on more than 60% of its cases, while 1/3 of the labs indicated such audit on less than 20% of cases), 23% indicated re-testing of samples, and 9% of the labs identified other forms of internal evaluation (See Figure III-17). In terms of internal audits involving re-testing of samples, the percent of cases re-examined varied widely between 0 to 100%. Half of the responding laboratories reported re-testing 6% or less of analyzed cases, while 3% of labs (n=2) indicated re-testing on 100% of their analyzed DNA cases.

Of those labs that indicated re-testing as a practiced form of internal review, 80% (n=11) indicated that approximately 100% of the re-tested results agreed with the original reports, while 20% did not specify the re-testing results. One lab stated that 99% of its re-tests agree with the original results because approximately 1% of the re-test results are outside the match window for RFLP analysis (+-2.5%), due to concentration problems. Furthermore, another laboratory suggested that disagreements between the re-test and the original result could be due to the mixtures and intensity of dots on DQA1 testings.

Figure III-17. Internal Review (n=66)



According to laboratories' description of their re-testing protocol, re-testing is either performed randomly or determined by the technical leader, reviewer, or the chief serologist. Furthermore, one laboratory stated that re-testing will likely be performed if the original analyst has left the laboratory. However, the most common factor influencing re-testing selection, as stated by laboratories, is whether or not there is enough remaining evidence to be re-analyzed.

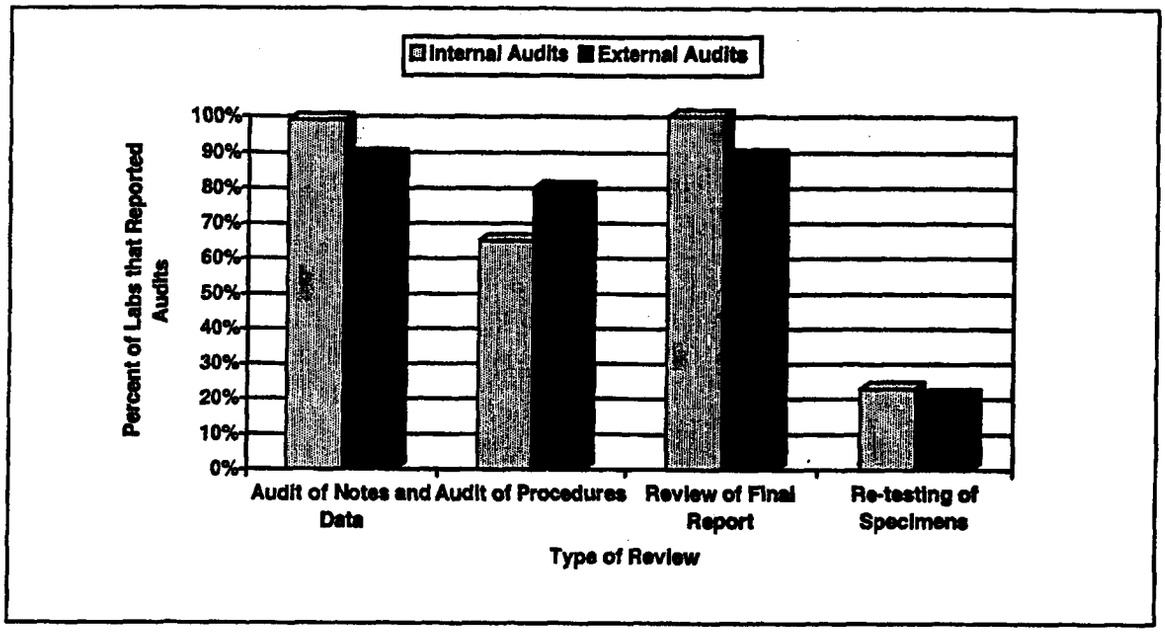
4. External Auditing / Re-Testing

Contrasted with laboratory's internal review procedure, the survey was also interested in scrutiny of lab work by personnel outside of the laboratory. According to responding labs, about 70% reported having external audits on DNA cases during 1997, with half of the laboratories reporting external audits on 10% or less of DNA cases analyzed. The average percentage of analyzed DNA cases receiving external audit per lab is 12%, and the median is 10%. Two labs reported that 100% of their cases were audited externally in 1997.

For the laboratories that experienced external auditing/review during 1997, when asked to elaborate on the nature of the external review, 89% of the labs indicated that the audits included review of notes and data, 80% indicated audits of procedures, 89% indicated reviews of final reports, and only 21% indicated re-testing of samples (See Figure III-18). According to the laboratories which had case samples re-tested, re-test results were in agreement with the original results 100% of the time. Of the labs that reported external audits, 13% indicated that their audit involved experts outside of the laboratory but from within the same state system, 58% of the labs indicated auditing of cases by experts outside the system but from other private/public labs, 25% indicated involvement of defense experts, and 30% indicated review of cases by external QA

Auditing Groups (NFSTC, ASCLD-LAB, and TWGDAM).

Figure III-18. Nature of Internal and External Audit



However, our findings show that the types of external audits also vary by the organizational location and the accreditation status of the laboratories. While majority of state and county labs received external audits by experts from private/public laboratories outside of their system, most city labs received external QA audits. In terms of variation by accreditation, accredited labs had a higher rate of external audit by private/public labs outside the system (65%), while 43% of non-accredited labs indicated such reviews - a decrease of 34%. Furthermore, fewer accredited labs reported external review by defense experts (18%), while 43% of non-accredited labs reported such audits – an increase of 38%.

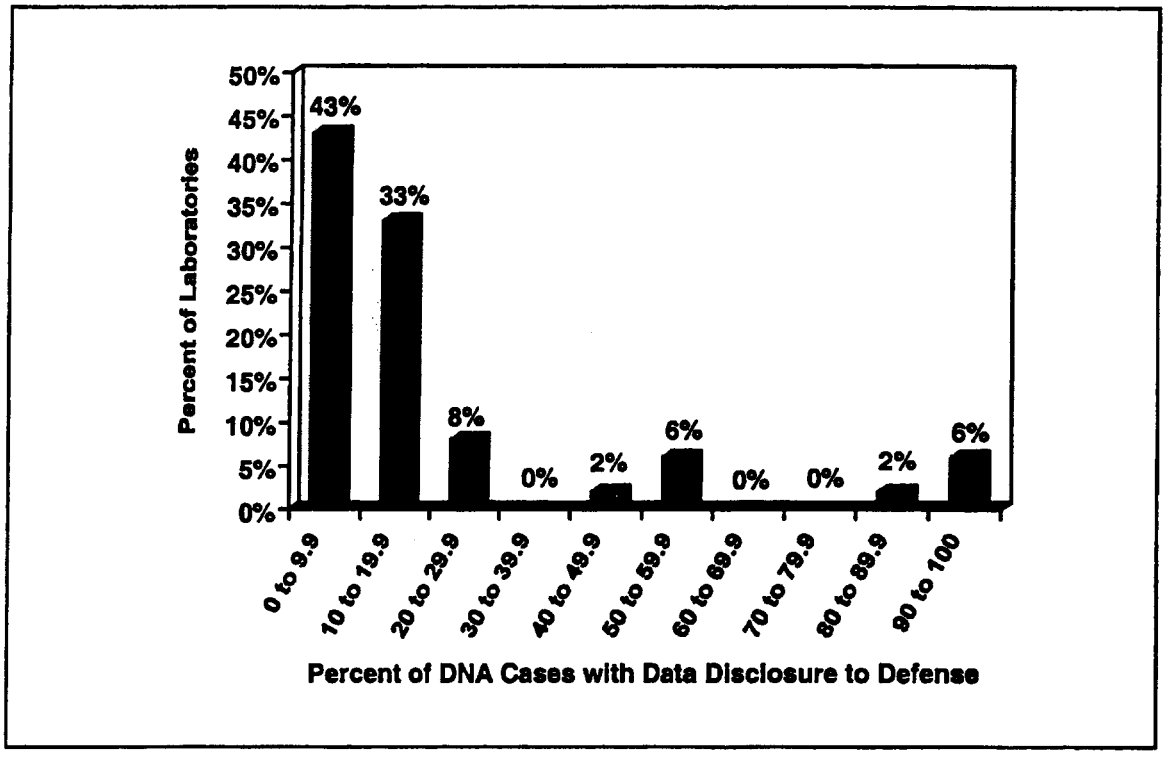
5. Defense Scrutiny / Discovery

When asked to estimate the percentage of DNA cases analyzed in calendar year 1997 for which an analyst (from respondent's laboratory) testified at a hearing or at a trial, the responses varied between 0 to 50%, with a mean of 15% of cases per lab, and a median of 11%. Half of the labs reported that less than 10% of their cases involved testimony, while one-third of labs indicated that more than 20% of their cases involved trial hearing testimony by analysts.
In terms of defense expert review of laboratory data, 75% of responding labs reported defense review on 10% or less of their DNA cases. While the median of the responses is 5%, one lab indicated that 100% of their cases involved defense expert review of laboratory report and/or other lab data.

Furthermore, laboratories were asked to estimate the percentage of DNA cases in which laboratory data beyond the lab report was disclosed to the defendant as a result of a discovery motion. (Laboratory data beyond the lab report includes laboratory notes, methods book, proficiency testing results, raw data, and other primary laboratory work products.) While half of the responding labs indicated disclosure on 10% or less of analyzed cases in 1997, a few labs reported disclosure on 100% of their cases to the defense (See Figure III-19).

In the interest of defense re-testing, 43% of the responding laboratories indicated zero defense expert re-testing on analyzed DNA cases in 1997. Of the 57% that reported some re-testing activities, the maximum percentage of cases involving re-testing by a defense expert, as reported by two labs, was 30%. When asked whether the defense expert re-tested case findings agreed or disagreed with original findings, 40% of responding labs either did not know, or were not informed of the results, while 58% of the labs reported that 100% of the re-test results agreed with original findings.

Figure III-19. Disclosure of Data to Defense (n=63)



C. Results of Survey of Defense Attorneys and Expert Witnesses—Phase 2

1. Results of Survey of Defense Attorneys

Of the 19 surveyed defense attorneys, 6 replied to the survey. In total, the respondents reported serving as the defense attorney of record for 56 cases involving DNA analyses in 1997. The survey instrument (see Appendix D5) asked for experience in the calendar year 1997, so that defense attorney, expert witness and lab survey results would all be for the same time period.

The reported median is seven cases, with a mean of nine cases per respondent. While three-quarters of the respondents reported less than 15 DNA cases, the maximum number of DNA cases served by a respondent was 25. Furthermore, respondents reported a total of 20 cases wherein they served as the legal DNA consultant instead of the defense attorney of record—two respondents reported 10 cases each, and the remaining respondents reported consulting on no DNA cases.

Because the number of returned surveys was small, it is not clear that the data are representative. The numbers and percentages in the following discussion should be read with this limitation clearly in mind. We interviewed some of the responding attorneys to try and get a better idea of what they perceived to be the major issues and problems with forensic DNA laboratories.

To understand the nature of defense evidence reviews, surveyed attorneys were asked to estimate the percentage of their cases that received various types of reviews. When asked about reviews of the DNA laboratory report, half of the respondents answered that more than 10% of their cases in 1997 consisted of lab report reviews only, while the other half indicated that none of their reviews consisted of only the lab report. According to the respondents, 83% indicated that half or more of their cases involved consultation with an independent expert about the laboratory report. While all respondents indicated that more than half of the time the underlying test results were reviewed, 83% of the respondents reported to have an independent expert review the underlying test results on 50% or more of the DNA cases. Also, 83% of the respondents reported that they visited the laboratories that performed the tests for half or more of their cases.

In terms of re-testing, half of the defense attorneys indicated that they had an independent lab either replicate or retest the same sample tested by the prosecution for 10% of their DNA cases, while one attorney reported such testing for 100% of the cases. Furthermore, half of the respondents denoted zero testing of additional samples (other than those tested by the prosecution), and the other half of the respondents indicated that they have an independent lab test additional samples for 10 to 25% of their DNA cases.

About 2/3 of the responding defense attorneys reported that their defense experts found problematic or questionable results and/or interpretations in the original lab work on 50% or less of their DNA cases. One-third of the defense attorneys reported that reviews on all their DNA cases resulted in problematic findings.

When asked to estimate the percentage of cases where the retest resulted in notable differences from the initial laboratory's results, half of the respondents indicated that they found notably different results in 50% or more of the cases they re-analyzed, while 1/3 of the defense attorneys found no notable difference in any of their retested cases. According to respondents, the nature of discrepancies in unmatched cases could be different interpretation of bands in STR testing and blue dots in PM/DQA1, or errors due to contamination and mixtures.

The respondents' experiences in obtaining laboratory data from the prosecution (e.g., laboratory reports/information, test data, QA/QC records, etc) varies between easy and very difficult. While one-third of the responding defense attorneys ranked the experience as "very difficult," the remaining respondents did not share a consensus on the level of difficulty of obtaining lab data. However, analysis shows that the respondents who ranked the experience most difficult also found 100% of their DNA cases to have problematic original lab results,

whereas the rest of the respondents who found the experience less difficult (77%) indicated that one-half or less of their cases had problematic original lab results.

In terms of securing funding for reviews and re-tests, the procedure varies among respondents. One-third of surveyed defense attorneys reflected positively on their attempts to secure funds for review of DNA evidence, re-tests, and expert witness fees, while one-third indicated that their own agency took care of such costs, so no extra petition was necessary. The remaining one-third of attorneys reported difficulties in securing funds for reviews and re-tests, and the main reason expressed was reluctance of the courts to authorize funding. One defense attorney stated, "Many times the DNA test is performed by the prosecution during the course of the case, while pending before court. Obtaining the continuance of the trial date to perform testing by the defense over the prosecution objection is the most difficult problem. Finding a lab within time constraints is also difficult." Furthermore, another problem that was brought up by attorneys was that court budgets vary by offense. According to one attorney, murder trials tend to have more funding available than rape cases, even although sexual offense cases employ more DNA tests of evidence than murder cases.

When asked about their general experiences, the majority of the respondents did not comment positively on DNA laboratory procedures. Most were offended by the secrecy surrounding notes, protocol, and raw data in testing, and furthermore, they questioned the validity of lab reports. False information, cover-ups, and improper interpretation of data were a few of the problems cited by the defense attorneys. Furthermore, the prosecution was also blamed for using tactics to prevent defense audits. However, one respondent indicated that lack of preparation by defense attorneys is also a problem, and suggested that the defense bar needs to have more experts to assist attorneys.

Interviews with several respondent defense attorneys did not add much to the observations discussed immediately above. There is a significant lack of trust of the forensic laboratories on the part of much of the defense bar. It may be unreasonable to suppose that it could be otherwise in such an adversarial system, but the situation that is created cannot be seen as very constructive.

2. Results of Survey of Experts / Expert Witnesses

The survey instrument (see Appendix D6) asked for a summary of the respondent's experience in calendar year 1997. The same year was specified in the Defense Attorney (Appendix D5) and the Phase 2 Forensic Laboratory (Appendix D2) surveys so that comparisons could perhaps be made later.

Of 27 surveys sent out to DNA expert witnesses, 11 surveys were returned (a return rate of 41%.) In total, the respondents had reviewed 238 DNA cases for criminal defense lawyers in 1997. Half of the respondents reported reviewing fewer than 10 cases in 1997, and 75% reported reviewing fewer than 20 cases. In addition, respondents reported a total of 53 DNA cases reviewed for someone other than a defense lawyer in 1997.

Because the number of returned surveys was small, it is not clear that the data are representative. The numbers and percentages in the following discussion should be read with this limitation clearly in mind. We interviewed some of the responding experts to try and get a better idea of what they perceived to be the major issues and problems with forensic DNA laboratories.

One objective of this element of the study was to get an idea about what problems were most commonly seen by these experts and whether there were identifiable trends in the perceived

problems. The experts surveyed in this part of the project generally review the work and results of the forensic lab that first examined the case. Another objective of this survey was to try to estimate what fraction of all forensic DNA cases are subject to any scrutiny outside the originating laboratory. Review by external experts could be viewed as a kind of blind proficiency testing, similar perhaps to external random audit / reanalysis.

(a) General Case Information

Two-thirds of the responding expert witnesses said that no retesting was performed in their own laboratory for cases that had previously been tested at another laboratory. The remaining 1/3, however, performed retesting on about half (55%) of the previously tested DNA cases. As for additional testing on evidence not previously tested, 2/3 of the experts indicated that no such testing was performed by their agencies during 1997, while 1/3 performed tests on previously untested evidence for less that 25% of their DNA cases.

Approximately half of the respondents indicated that they had referred a previously tested sample to another laboratory (other than their own) for a retesting. While most experts referred less than 30% of their cases to another lab, one expert reported having referred 100% of his cases to another laboratory for retesting.

As for referrals on evidence not previously tested, 1/3 of the respondents indicated that they had referred an item not previously tested to another laboratory for DNA testing. Two of these respondents referred less than 20% of their cases for external testing on previously untested evidence, the other two referred 90% or more of their cases for such testing.

Although 80% of the respondents indicated that a fraction of their 1997 DNA cases contained problematic results and/or interpretations made by the original laboratories, the actual percentage of questionable DNA cases varied among experts. While the average percentage of problematic DNA cases is 25% per expert surveyed, ¼ of the respondents reported that they found problems in less than 7% of their cases, and another ¼ of the respondents reported errors in more than 75% of their cases.

In attempting to mitigate the likelihood of problematic results, respondents raised the possibility of "test observation" as an important tool useful for catching laboratory errors or problems in testing. In some cases, court orders may be obtained to allow a defense expert to observe the DNA testing in the primary laboratory.

Expert witnesses were also asked whether or not they had access to materials from other laboratories that were necessary to perform an adequate review of DNA test results. While 20% (n=2) responded that they have always had access to materials needed, 80% of responding experts indicated that they did not. According to the latter respondents, requested materials that were not provided to them included:

a) Original or excellent reproductions of autorads, gels, quantiblots, and/or typing strips,
b) Details of PT reports, and details of certain experimental procedures, and
c) Detailed lab notes and high quality photos of original results.

When defense experts elaborated on reasons given for declining their requests, time and price were brought up as possible causes. However, the majority of respondents focused on two main reasons as to why they thought their requests were refused: 1) policy and legal limitations forbidding distribution of certain information; and 2) poor quality of laboratory records and /or work product. These respondents were suggesting that, because laboratories do not always document adequately, they are reluctant to give out detailed documentation for fear of criticism.

(b) Nature of Problems Found in Cases

To clarify the nature of problems that expert witnesses encounter, the survey instrument also asked the respondents to point out specific problems and estimate the percentage of reviewed cases in which these various problems occurred.

Chain of Custody: Chain of custody was cited as an important aspect of DNA analysis because problems of this type are often hard to detect (e.g., switching of samples and poor documentation of sample identity). Survey findings indicate that 40% of expert witnesses found chain of custody problems to be an issue in an average of 14% of cases, while the remaining 60% of experts indicated that they did not find chain of custody problems in their reviewed cases. As an example, one respondent pointed out that photographic documentation of PCR tests strips is often inadequate, thus making independent confirmation of lab tests impossible. Furthermore, related problems such as samples without labels and improper substitution of cases were also reported by respondents.

Sample Handling: Sixty percent of expert witnesses found sample handling issues (errors leading to cross-contamination) within their cases. Approximately half of the respondents found the problem in more than 12% of their DNA cases. One expert witness reported that in one case, evidence items were packaged together in a single container, causing sample contamination when PCR typing had to be done.

Inadequate Documentation of Procedures: While documentation of procedures varies from one laboratory to another, 80% of expert witnesses found inadequate documentation of laboratory procedures within their reviewed cases. While approximately half of the respondents reported that less than 25% of their DNA cases contained documentation error, on average, experts reported that they found inadequate documentation in 40% of DNA cases. Problems mentioned by respondents included:

a) Inadequate documentation of the quantity of DNA obtained from evidence samples;
b) Inadequate documentation of microscopic procedure (common problem with PCR-based tests); and
c) PCR-based tests performed without first testing for indications of relevant biological material.

Furthermore, one respondent reported that for some cases only the conclusions of the testing were stated, but no results listed in the report.

Failure to Follow Protocol: Half of the responding expert witnesses reported that they found cases in which the original analysts failed to follow the proper protocol for testing (for example, procedures as determined for HLA-DQA1, D1S80, and PM tests.) Although the estimated percentage of cases containing such error varied widely among experts, respondents estimated that they encountered this problem in 25% of DNA cases on average.

Poor Laboratory Practices: Seventy percent of respondents found what they considered to be poor laboratory practices (e.g., failure to account for failed controls or failure to perform controls) within their DNA cases. Furthermore, respondents also reported questionable issues with police evidence collection procedures, inadequate transportation, and improper storage conditions.

Biased or Problematic Interpretation: Seventy percent of respondents found that a fraction of their DNA cases contained biased or problematic interpretation of results by the original lab. Although interpretation is a difficult matter to evaluate, problems indicated by surveyed experts included:

a) Improper evaluation of weight of possible lab error;
b) Subjective interpretation of band intensities (mostly in HLA-DQA1 & PM tests); and
c) Misinterpretation due to incorrect statistical approach or inadequate protocol.

While half of the respondents indicated finding the error in less than 20% of their cases, ¼ of the respondents indicated that more than 50% of their cases contained such error.

Calculation of Inclusion Probabilities/Population Genetics Issues: Half of the expert witnesses found that a fraction of their reviewed DNA cases contained problems related to calculations of inclusion probabilities / population genetics issues. The mean percentage of cases thought to contain such error is 32%, with a median of 13%. Specifically, the problems indicated by responding experts included:

a) Overlooked statistical implications of potential suspects who were genetically related to the prime suspect;
b) Improper statistical evaluation of database searches or cases where large numbers of suspects were screened; and
c) Excessively precise random match probabilities, which gave the illusion that the estimate was well based.

Other Errors: Two expert witnesses found other types of error within their DNA cases, including: a) failure to resolve discrepancies between conventional serology typing and DNA testing; b) problems with the preparation or labeling of reference blood samples; and c) problems with the drawing, or the documentation of drawing, of reference blood samples.

(c) Evidence Testing / Re-Testing

After performing retests on previously tested evidence, 80% of responding expert witnesses found no significant difference between the retest results and the initial laboratory's results. One expert witness (the remaining 20%) found disagreement in about 5% of the analyzed cases.

According to one expert witness, the original analysis is rarely so faulty that the results are completely wrong. Rather, mistakes made by prosecuting agencies tend to occur *before* the DNA lab gets the evidence (collection and preservation), or *after* the analysis is completed (interpretation). While stating that retesting is a useful tool, the respondent suggested that the main issue needing to be addressed is education -- for individuals responsible for collecting evidence at the scene, for police officers, for crime lab personnel, and for the analysts responsible for interpretation.

In the cases involving testing of additional items not previously tested, 2/3 of responding experts found that such additional testing produced important or unexpected results (e.g., indication of additional or alternative perpetrators). Most respondents agreed that findings from additional testing often implicated third individuals, and in one specific case, additional testing on an item previously considered "inadequate" by the original lab led to a homicide conviction. Although the percentage of cases with significant results differs between 20 to 90% for respondents, expert respondents indicated that in an average of 34% of cases that involved additional testing, significant results were obtained.

Approximately 80% of responding experts found that in some cases, critical biological evidence was consumed in its entirety by the original laboratory in its initial DNA testing. On average, about one-quarter of cases were found to involve evidence that have been completely consumed. However, respondents indicated that the percentage of cases involving total consumption of samples has recently decreased with use of PCR typing and changes in laboratory protocols requiring labs to save portions of the evidence for re-test or additional testing.

(d) Results of Interviews

As noted above in § III.C.2, a few respondents were interviewed to see whether any further insight could be gained beyond the questionnaire response data. We tended to try to focus on how frequently the respondent thought there were significant problems, problems that might be considered serious enough to have an effect on case outcomes. The responses were mixed. Half the respondents tended to report that they almost always saw serious problems in detailed case reviews. Serious problems included things like tests where controls failed but the results of which were reported anyway, significant errors in calculations, and reporting or ignoring (depending on the point of view in the case) "subthreshold" signals. The most frequently mentioned problem was simple failure to follow the laboratory's own standard protocols and QA/QC guidelines. The other half of the respondents indicated that, while problems of one sort or another could always be identified, serious errors were seen in 20 to 25 percent of the cases reviewed. One respondent had reviewed several hundred cases over perhaps a decade, and believed that these could in some way be considered a random sample.

Some of the problems identified did not involve the laboratory or its personnel, but had to do with evidence collection, packaging or storage. The experts interviewed recognized that these issues were not laboratory problems.

There is a sense among reviewing experts that analytical results per se are generally fine, and that they are generally reproducible in another laboratory. Problems that occur are the result of actions before the evidence gets to the laboratory, and/or in the interpretation of the results obtained. There is also a sense among the reviewing experts that it is never easy to get data or information about a case one is reviewing. They tend to place the blame for this more on police administrators and/or prosecutors than on laboratories, though it is noted that the outcome is not changed.

3. Conclusion from Phase 2 Surveys

In Phase 2, surveys of forensic DNA laboratories focused on their sample retention practices, and internal and external reviews of casework. Evidence retention, of course, is a necessity if that evidence is to be re-examined as a quality assurance mechanism. Our findings showed that more than 90% of laboratories retain the original specimen, extracted DNA, or both,

and the greater majority of labs (73%) reported only about an average of 5% of their DNA analyses involved total consumption of key biological samples. Almost all of the responding labs (99%) estimated that between 95% and 100% of their DNA cases were subjected to internal reviews, with the most common type of reviews being a review of final report. The percentage of analyzed DNA cases receiving external audits is lower than internal reviews, with the most common type of external audits being a review of the final report. Rates of re-testing for both internal and external audits are quite low, with most of the re-test results agreeing with the original reports.

Six defense attorneys and 11 expert witnesses also returned surveys in Phase 2. Unlike the labs that indicated that a large proportion of the re-test results were in agreement with the original analyses, half of the responding defense attorneys indicated that they found discrepancies between re-test results and original reports. The majority of the responding defense attorneys did not comment positively on DNA laboratory procedures, many questioned the validity of lab reports, and one respondent also pointed out the need for defense attorneys to have greater assistance from technical experts.

According to independent expert witnesses, rates of re-testing of cases were very low. Only one-third of the 11 responding expert witnesses indicated that they performed re-tests on DNA cases, and of these, experts performed re-tests on only about half of the cases they reviewed. Many expert witnesses suggested that they didn't have access to materials and documentation from the primary testing labs that were necessary to perform adequate re-tests. The experts believe that due to the fear of criticism from expert witnesses, laboratories are sometimes unwilling to give them detailed documentation. They reported, in some cases, critical biological evidence was totally consumed, so a re-analysis was impossible. However, they also pointed out that because of the growing use of PCR typing and changes in laboratory protocols requiring labs to save portions of the evidence to allow for re-analysis, the fraction of DNA cases where key samples are totally consumed is decreasing.

One of the objectives of Phase 2 of the project was to try to see whether there was already significant external review of DNA cases, by an external entity engaged by the primary lab or by the defense and its experts. The data show that there is some level of external review, but it probably does not cover a very large percentage of the cases worked. Non-adversarial external reviews tend to be limited to audit and record review, and rarely seem to involve retesting. Some experts who review cases indicate that there are serious problems in almost every case they see, where others suggest that serious problems are seen in a minority (perhaps 20 to 25%) of cases. A similar picture emerges from the survey of defense attorneys themselves. Part of the reason for trying to determine the extent of current external review of cases was that it could be regarded as an external blind proficiency test of a sort.

Another objective of Phase 2 was to determine to what extent biological evidence might be preserved, and thus available for random audit / reanalysis. The data suggest that a majority of labs retain cuttings and/or extracted DNA for potential retesting. While the fraction of cases that are externally scrutinized is probably relatively small, it is also true that the number of external blind PTs that could be administered to labs in a given year is also quite small, given the costs and complexities.

For the majority of laboratories, there is probably enough extracted DNA and/or stain cutting retained in casework to make possible a relatively random retesting program. It is noteworthy, however, that such retesting would serve primarily to check analytical results. Most

labs do not retain entire items of case evidence or all the items in a case. It would be more difficult, therefore, to "audit" the laboratory's handling of the case, apart from its analytical results.

IV. BLIND PROFICIENCY TEST FEASIBILITY TRIALS

A. Phase 1 Feasibility Trials

In this project, we decided to direct actual blind trial proficiency tests to laboratories that had agreed in advance to be potential test candidates (or reference laboratories). In this section (§ IVA), the Phase 1 feasibility trial tests are discussed. Phase 1 tests consisted of ten blind case set-ups for submittal to DNA laboratories (two of these cases were submitted to the FBI laboratory by its own blind PT contractor who worked with us), and two reference laboratories. The Phase 1 tests were manufactured and administered in two cycles, five tests at a time. Phase 2 tests are discussed in a separate section, § IVB, below.

Given the nature of this testing and the fact that these laboratories and DNA analysts could be categorized as "research subjects", we were ethically bound to secure the permission of these laboratories to engage in this testing before it commenced. Our survey of every laboratory that we determined to be performing forensic DNA analysis resulted in the return of 94 surveys (see § IIIA), 38 of which agreed to be potential participants. Another laboratory from which survey data was not obtained agreed to participate in the actual feasibility trials. Thus, the number of potential participating labs from our survey was 38, but in fact, there were a total of 39.

The thirty eight potential target or reference laboratories (from our survey) had the characteristics shown in Table IV-1.

Table IV-1. Characteristics of Target and Reference Laboratories – Phase 1

Number (%Total)	Type	Service Area	No. DNA Analysts (Avg)**	Cases Come Mainly From
17 (44.7)	State System	Entire State*	n = 17 (7.35)	Police / Other Labs
10 (26.3)	County	Entire County	n = 8 (2.75)	Police
5 (13.2)	City	Entire City	n = 5 (5.6)	Police
6 (15.8)	Private / Independent	No Limits	n = 6 (4.67)	Police, Attorneys, Other Labs

* not necessarily the sole provider of DNA typing services in the state
** n = number of labs who reported the number of DNA analysts

1. Selection of and Agreements with Participating Laboratories

Laboratories that agreed to be potential participants did so by signing a formal agreement with us (Appendix E1), containing certain mutual understandings. We assured participating labs that: (i) specimens in the fictitious cases would be manufactured following the "Guidelines for DNA Proficiency Test Manufacturing and Reporting" [Crime Lab. Digest 21(2,Apr):27-32, 1994]; (ii) we would notify the lab immediately that the case in question was a blind proficiency test, once DNA testing and reporting were completed, and we would let them know how the results compared with those of the reference laboratories; (iii) we would not reveal the identities of participating laboratories in our project materials/reports nor the names of specific laboratories or specific examiners actually participating in any trial testing in this project unless we were legally required to do so; and (iv) participation in this project was totally voluntary, and could be

discontinued at any time.

On their part, the participating labs agreed that: (i) they would contact the project office if they suspected a submitted case was really a "blind" proficiency test (and we would tell them if their suspicion was correct); (ii) they would keep confidential their involvement as a trial test site and/or as a reference laboratory, unless they were legally required to reveal their participation; (iii) they would either purge from their files (including computer-stored records) all records connected to a trial blind proficiency test "case" introduced into the lab, or clearly identify the records retained as being connected to a blind proficiency test; (iv) they would completely purge from their databases, and cause to have purged from any centrally-maintained databases, any and all DNA types and profiles that were entered into such databases as the result of analysis in a "case" that we revealed to them was fictitious; and (v) they would cooperate with us in analyzing the results and problems, and assessing the costs, of conducting blind proficiency tests.

Eight of the 39 potential target laboratories were selected for testing in this project. The eight were selected on the basis of being: (i) reasonably representative of the different types of labs; (ii) spread over a large area and range of jurisdictions; (iii) accessible through a conduit laboratory, or law enforcement agency, that would cooperate with us in submitting a blind proficiency test case. Some of the 39 labs that agreed to be potential targets for blind proficiency tests were not accessible through their usual submitting agencies for purposes of introducing blind proficiency test cases in this project, because those agencies would not agree to participate. Thus, the actual number of potential participating laboratories was fewer than 39. More details about the characteristics of the labs actually tested are given below.

The FBI Laboratory (listed as "federal" in the laboratory characteristics table below) made arrangements with a separate contractor, who worked with us in manufacturing and arranging to submit cases, to submit cases to them. In what is called Phase 1 of this project, two cases were submitted by that contractor to the FBI Lab, one in each cycle of Phase 1. The FBI Lab was required by the DNA Act to subject itself to blind proficiency tests.

It may be of interest to note that some of the reasons given by law enforcement agencies for their refusal to participate in blind testing included: (i) the mock case paperwork that would be required to support submission of a blind proficiency test case would constitute a violation of police departmental procedures or statutory prohibitions against filing knowingly false official statements; (ii) participation would involve knowing deception of the lab personnel by police officers and would constitute a breach of the trust between the law enforcement agency and the laboratory; and (iii) the agency had full confidence in the laboratory, and could not see any compelling reason for helping to engage the laboratory in blind proficiency testing.

2. Blind Trial Proficiency Test Case Setup, Manufacture and Distribution

We entered into a subcontract with a TWGDAM-approved DNA proficiency test manufacturer under the terms of which the contractor would manufacture to our exact specifications the proficiency tests designed by the UIC project and the two proficiency tests designed in collaboration with the independent contractor tasked with submitting blind DNA proficiency tests to the FBI Laboratory.

Approval was obtained from the UIC Institutional Review Board (IRB) for the employment of volunteer human subjects as donors of biological evidence specimens, and approval was obtained as to the substance and form of the informed consent statement. The informed consent statement was provided to the manufacturing contractor for use with the volunteer donors

(Appendix F). The specimens obtained from the donors were anonymized by the contractor, and the signed consent returned, sealed, to one of the UIC project directors. Knowledge of the identity of the volunteer donors was thus restricted in effect to one UIC project director. The informed consent statement that was devised and approved by the IRB, explicitly informed the volunteer donors that their DNA profiles could be databased for some period of time until the "case" was revealed as a blind proficiency test to the target DNA lab. There was a very small, but non-zero potential for a volunteer donor in this project to experience difficulties if some life event involved that person with a real case involving biological evidence during the time the individual's profile was databased because of this project. We were subsequently advised by NIJ General Counsel that the identities of our volunteer donors in this project are protected under any circumstance or legal proceeding.

In preparation for the actual manufacture of "biological evidence," we recruited several male and female volunteers to purchase complete sets of clothing, then take the clothing through a specified set of wear-wash-wear cycles. The identities of these individuals were recorded in the event that specimens for DNA typing might later be required. In addition, we agreed with our manufacturing contractor that some volunteer biological-specimen donors would contribute their own well-worn clothing for use as "evidence" in some of the "cases."

Prior to actual manufacturing of biological evidence, "cases" were set up with knowledgeable people in law enforcement agencies or labs who would serve as submitters—conduit labs (see below). This step involved interviews to determine the detailed characteristics of a "typical" case, what type(s) of evidence would typically be submitted, how it would be packaged, how the packaging would be marked and by whom, what paperwork would accompany it, etc. To minimize the number of specimens—and thus the chances of introducing an item or items that would result in detection of the "case" as a proficiency test, part of the setup involved determining the minimum number of specimens that the lab would normally expect to receive in the type of case selected. A typical setup involved a crime scenario, data from which was used by the submitter in completing lab submission forms, and by the manufacturer in putting together the biological evidence items. Suspects and victims were assigned fictitious names, dates of birth, and sometimes race, and an offense date that made sense in terms of the normal submission lag time was selected.

The detailed setup agreement with the submitter, and the case scenario, then enabled us to prepare detailed instructions for the manufacturer to put the biological evidence together. Complete specifications for a "case" included detailed instructions for the collection of specimens from donors, for use of those specimens in manufacturing the biological evidence, and for transmitting the items to the submitting entity. We also included a document to be sent to the submitting law enforcement agency or conduit lab along with the "case." In some instances, additional instructions for marking the evidence, etc., was included with this document. A typical example of these detailed instructions and specifications is given in Appendix G.

A total of five volunteer donor individuals (three females and two males) contributed biological specimens to these tests. As "cases" were constructed from scenarios prior to manufacturing, detailed requirements were written for the nature and quantity of biological specimens required from each donor. Some female donors contributed panties worn for at least one day and not laundered prior to submission. Semen was counted for sperm density before spiking swabs and/or panties. Under the assumption of 2.5 pg DNA/cell, swabs and/or panties

were spiked with semen sufficient to contain at least 1-2 μg sperm DNA, more than enough in theory for RFLP typing.

Agreements were made and signed with law enforcement agencies and laboratories (conduit labs) that participated as submitters in this project. A typical agreement with a law enforcement agency is shown in Appendix E2, and a typical agreement with a conduit lab in Appendix E3. These agreements are very similar to one another. As in the agreements made with the potential target laboratories, these agreements contained certain mutual assurances.

Blood specimens from the donors were forwarded to two reference laboratories for DNA typing at about the same time the "case" specimens were forwarded to submitters. In addition to blood specimens, "copies" of manufactured biological evidence items were sent to the reference labs as well. By "copies" is meant contemporaneously manufactured in replicate as closely as possible. In this way, the reference labs were able to conduct some of the same procedures on items like semen-spiked vaginal swabs or semen-spiked panties as the target labs. Such reference lab data can be helpful in diagnosing any problems with the manufactured "evidence," and also in interpreting a target lab's results (e.g., a target lab may report "insufficient DNA for RFLP typing," and reference labs may both obtain quantities of DNA well in excess of that normally required for RFLP analysis). In addition to potentially assisting in the interpretation of proficiency test results from target labs, furnishing two or more specimens from the same individuals provides an opportunity for replicate typing and/or RFLP band sizing by the reference labs.

The manufacturing contractor maintained detailed and meticulous records of the specimen collection, testing and use in the manufacture of blind proficiency test cases. This record keeping constituted a significant part of the contractor's QA procedures for manufacturing the specimens and cases. To avoid insofar as possible any possibility of specimen mixups or other errors in case manufacturing, only five cases (and their replicate specimens) were manufactured at the same time. Samples of all the biological specimens used in case manufacturing were retained during the project period as well in case any question about a specimen arose in the future. We estimate that a maximum of ten blind cases could be manufactured simultaneously following rigorous the rigorous QA procedures that were followed in this project. This estimate has implications for the cost estimates provided later in this report for a larger-scale blind proficiency test program.

In this project, blind proficiency test cases were introduced to DNA laboratories in one of two ways: through law enforcement agencies; and through other laboratories who did not do DNA testing themselves, and who regularly submitted evidence for DNA typing to the target lab. Thus, for this project, a "target" lab is a DNA typing lab that actually received a "case" for DNA typing. A "conduit" lab was another forensic laboratory that does not do DNA typing, but regularly sends cases or evidence items to the target lab for DNA typing. Some conduit labs were forensic labs that regularly submit evidence to private, independent labs for DNA typing. Others were part of state systems where DNA typing is restricted to a central location or to a few labs in the system. We made contact with appropriate conduit laboratories just as was done with law enforcement agencies. As noted above, an agreement was signed with the conduit labs that was very similar to that used with law enforcement agencies submitting "cases" for this project.

3. Modalities of Blind Proficiency Test Introduction and Their Associated "Levels of Blindness"

We recognized early in the project that there are several different modalities of introducing blind tests into DNA laboratories. Different "levels of blindness" are associated with these

models. possible in blind proficiency tests. There are four potential models for blind proficiency test introduction, described below and discussed more fully in §V.

1. Blind / LE - Blind to everyone in the target lab; submitted via law enforcement agency (LE)
2. Blind / CL - Blind to everyone in the target lab; submitted via another lab that does not do DNA typing itself – the conduit lab (CL)
3. Blind Analyst - Only the analyst is blinded
4. Random Reanalysis

The first two of these might be defined as "fully blind," in the sense that no one in the target lab is informed about the test or even the possibility of a test in advance. The Blind / LE is the most challenging and probably most expensive of all the models for the testing organization (see below at § IVA.6). In our limited studies, both the "fully blind" models succeeded. One Blind / LE test "failed," in the sense that it was detected by an evaluating criminalist, but it went forward anyway as a "Blind Analyst" test (see below at §IVA.4).

For the purposes of the feasibility segment of this project, blind proficiency tests were realistic, fictitious cases introduced to a target laboratory either through an external law enforcement agency or a conduit laboratory. Although, as noted above, the labs had agreed in advance to be potential targets because of the nature of this project, none of the personnel in target laboratories knew in advance whether they would receive tests. And, of those that were tested, no one was informed about any tests until they had been completed.

4. Results

(a) Details of the Actual Blind Trial Feasibility Tests

Eight separate blind proficiency tests were carried out during Phase 1 of the grant period. We also worked in close collaboration with the contractor responsible for submitting blind DNA proficiency tests to the FBI Laboratory. Two blind tests were submitted to the FBI Lab. Thus, taking into account the collaboration, ten blind tests were set up, constructed and submitted. A total of five volunteer donor individuals contributed biological specimens to these first ten tests. Fresh whole blood specimens from each person, anticoagulated in disodium EDTA, were forwarded to two separate reference laboratories for complete DNA typing. In addition, the vaginal swabs spiked with seminal fluid that were used for blind proficiency test "cases" were replicate manufactured as closely as possible and forwarded to the reference laboratories, in some cases along with oral and/or anal swabs from a female donor.

The characteristics of the ten Phase 1 blind proficiency tests are shown in Table IV-2.

Table IV-2. Characteristics of the ten Phase 1 Blind Proficiency Tests

Test	Target Lab Type	Submission through[a]	Type of Case[b]	Type DNA Testing Done[c]	Reported Findings	Turnaround Time (m)[d]
1	Private	CL	SA	PCR	Suspect included	0.83
2	Private	CL	SA	PCR, STR	Suspect included	1.63
3	State	CL	BT	RFLP	Blood on suspect clothing consistent with victim	5.06
4	State	LEA	SA	RFLP	Suspect included	6.9
5	Federal	LEA	SA	RFLP	Suspect included	16.53[e]
6	State	LEA	SA	PCR	Suspect included	3.46
7	Municipal	LEA	SA	RFLP	Suspect included	1.00
8	State	LEA	BT	N/A	Not completed[f]	3
9	State	LEA	SA	RFLP	Suspect included	4.06
10	Federal	CL	SA	RFLP	Suspect included	6.36

Table Footnotes:
[a] CL: conduit laboratory; LEA: law enforcement agency
[b] SA: sexual assault; BT: blood transfer
[c] RFLP; PCR: HLA-DQA1 and/or PM and/or D1S80; STR; N/A: not applicable
[d] Months, obtained by consistently dividing turnaround time in days by 30. The turnaround time is calculated from the date of submission of the last specimen in the case to the date of the laboratory's report.
[e] Hairs and fibers unit completed its report with a turnaround time of 2.16 m
[f] Lab request to police for additional specimen not communicated in a timely way to project team; see in § IVA.4.(b) below

There were 2 tests to private, for-profit labs, 5 tests to state labs or labs that are part of state systems, 1 test to a municipal (city) lab, and 2 tests to the FBI Lab. The tests were manufactured and submitted to forensic DNA laboratories in two cycles, each phase consisting of five "cases" that were contemporaneously manufactured. Cycle 1 manufacture was completed around August 6, 1997. Three UIC cases and the FBI case were submitted within a few days. One UIC case submission was delayed for several weeks. Cycle 2 manufacture was completed around October 20, 1997. All those cases were submitted within a few days.

Four tests went to DNA labs via "conduit" forensic science labs that do not themselves do DNA typing. Of those, one of the "conduit" labs was part of the same system as the DNA lab. Six others went to DNA labs via law enforcement agencies. Of those, two were submitted by the police agency that is the laboratory's parent organization. One case was submitted by a law enforcement agency to a lab that is part of a state system; that lab did the initial work-up, then forwarded the biological evidence to another lab in the system where the DNA typing was done.

(b) Detection / Communications Failures in the Blind Test Feasibility Trials

One of the blind proficiency test cases was detected by the target lab. It was a sexual assault case—primarily a sexual assault evidence collection kit—submitted by a law enforcement agency to the lab. The criminalist screening the case noted slight discrepancies in the time recorded on the packaging for the collection of some items in the kit. The criminalist also noticed that the vaginal smears (slides) were not streaked in the manner typical to the jurisdiction. In this instance, we had not actually discussed the detailed preparation of slides with hospital personnel in the jurisdiction, and we did not do a detailed review of the markings on the packaging prior to case submission. In this case, the supervisor sent the case on to DNA typing without further comment. The net result was that the DNA analysts did not know that the case was a blind proficiency test until the testing had been completed.

The discrepancy in the way the slide was made points to the necessity for extraordinary attention to small details in blind proficiency test manufacturing, especially in jurisdictions where there may only be a few clinical people actually involved in collecting sexual assault or other evidence from crime victims. Ensuring that such details are accurately duplicated in blind proficiency test case preparation necessarily increases the costs of this type of testing, as every jurisdiction has somewhat different practices. Manufacturers and test administrators cannot rely solely on law enforcement personnel to know all of these details. The discrepancy in the collection times for certain items in the blind test that was detected resulted from a failure of test administrators to review or supervise the marking of the evidence-collection kit item packages before submission. This level of detail in review and supervision of case submission would in many instances require one or more visits to the target lab sites by test administrators.

In another blind trial test, the case was successfully submitted to the laboratory. The case scenario involved a victim having injured a suspect during an assault, such that some suspect blood got on the pocket of his own clothing. During the struggle, the pocket bearing the blood was forcibly torn from his shirt. The victim then gave it to the police. The bloodstained pocket was submitted first, followed about a month later by the suspect's known blood, under a scenario where it took the police a while to locate the suspect. About three months later, the lab got to the case, and requested a specimen of victim's blood for comparison. The police tried to argue around the request—unsuccessfully—but did not communicate the request to the project team. Under these circumstances, the lab did not work the case because it went against their normal policies and practices. This test was ultimately terminated, because many months elapsed before we became aware of the lab's request for the additional specimen. This failure in communication between the project staff and the law enforcement agency was apparently the result of our failing to make sufficiently clear to the police personnel handling the case that any communications from the laboratory should be discussed with us immediately. Had that discussion occurred, the additional specimen could have been provided for submission. Although it is not specifically a test failure, one of the cases submitted by the FBI Lab's contractor to the FBI Lab was never completed.

(c) DNA Typing Results

All the labs reported the correct results, in the sense of including or excluding a possible depositor of biological evidence. The sexual assault cases were all "suspect included" cases. One blood transfer case involved victim's blood on a suspect's shirt, and the other blood transfer case

involved a suspect's blood on the pocket of his shirt, where the pocket was torn away during an assault and given to the police by the victim.

Some target labs did RFLP, some did the PCR-based loci HLA-DQA1, PM and/or D1S80, and one did several STR loci.

Reference labs typed six or seven RFLP loci, HLA-DQA1, PM loci, D1S80, and one did a number of STR loci. The reference lab data along with the target lab data is collected together in Appendix H1. Some target labs did not report RFLP locus band sizes, while others did. One target lab typed RFLP loci that were not typed by the reference labs, so no comparisons were possible. Reference lab data was shared with every target lab following the conclusion of the blind tests.

One target lab found that there was insufficient human DNA for RFLP typing. Following discussion with the submitter, the lab did PCR-based locus typing. Both reference labs in that instance, however, recovered more than sufficient DNA for RFLP typing from the duplicate specimens that had been contemporaneously manufactured.

(d) Turnaround Times

Generally, as can be seen in the table above, turnaround was significantly slower in government labs as compared with the private labs. The fastest turnaround was 25 days, and the slowest cases that were ever completed took 16 and 17 months.

In one case, a private lab that typed the case specimens using PCR-based procedures because there was insufficient DNA for RFLP typing was asked nevertheless if it would type the suspect known specimen by RFLP for possible CODIS query. The lab agreed to do this typing, but the turnaround time for the RFLP typing of the suspect's known specimen was significantly longer than that for the PCR-based typing—by a factor of about four.

(e) CODIS Issues

Two pairs of the actual blind proficiency tests had potential case-to-case, cross jurisdiction CODIS matches built in. The same male was the depositor on each of the case pairs, and the "suspect" in each pair of cases had the same surname. At the time of manufacturing and submission of these cases, it was not clear to us whether the NDIS component of the CODIS system would be fully operational.

One of the potential CODIS matches built into the tests was never found. The first reason for this situation is that the laboratory that received one member of the pair was not connected to CODIS at the time the case was completed. The second reason is that the first laboratory was asked to purge the blind test "suspect" DNA profile, and it did so, long before the 2nd lab ever worked their case. The second potential CODIS match was found.

A case submitted and worked in the Midwest was databased; later, the other case in that pair (a case submitted by a state lab to the FBI Lab for DNA typing) was worked and the match was detected. We had already notified the Midwestern lab that their "case" was a blind test, so they could immediately inform the FBI Lab of the situation.

(f) An Independently Executed Blind Proficiency Test

Independent of this project, a blind proficiency test case was successfully introduced into a laboratory in one of the states by an oversight group. This test—a sexual assault case—was not

detected by the target laboratory, and to our knowledge, the laboratory's results coincided with those of a reference laboratory. There was some discussion about further collaboration between this project and that jurisdiction in manufacturing and submitting additional blind tests, but nothing came of those discussions.

5. Fate of Records Created as the Result of Blind Tests

Purging of documents at the close of the research project was discussed with the Advisory Committee. While the project staff initially believed that every effort should be made to purge all records and data, the panel members advised that it might not be possible to purge all written records, and efforts to remove proficiency records could conceivably compromise the agency/laboratory's chain of custody/record keeping system. In fact, maintaining the records of all external proficiency tests could even be desirable since they document the laboratory's involvement in such programs.

The advisory panel did agree, however, that any DNA typing profile information from a blind test entered into the computer database had to be expunged to protect the project's anonymous donors. As part of our post-test notification procedures, we included a form (Appendix I3) on which the target lab administration certified to the project directors that any DNA profiles from the blind test which had been databased had been purged.

6. Different Introduction Modalities / Types of Blind Proficiency Test Cases / CODIS

(a) Introduction Modality

We predicted, and found, that the manufacturing and submission of blind proficiency tests through conduit laboratories is simpler than submittal through a law enforcement agency. A major reason is that conduit laboratories typically work the cases to some extent before deciding to send biological-evidence items off for DNA typing. In the Blind / CL model, it is fair to say that the conduit labs do the "criminalistics" part of the case, while the DNA labs determine the DNA types for the selected evidence items. Thus, to manufacture a blind proficiency test "case" of this kind, fewer specimens are typically required, less paperwork is expected to accompany the specimens, and the target DNA labs are further removed from the case facts, from the investigators, and are less likely to become suspicious. Some DNA labs request or require a summary of preliminary or presumptive test results and/or classical genetic-marker typing results. In a blind proficiency test, those data can be made up and supplied, since the DNA lab is not going to repeat the initial tests. From a manufacturing viewpoint, only biological-evidence items need to be manufactured. The DNA labs do not expect to receive entire items of clothing—cuttings will generally suffice—nor intact rape kits. Reference bloods can generally be submitted as dried stains on blot cards or on cloth.

Law-enforcement agency submissions are more complex for the blind testers in several ways. "Cases" from law enforcement to DNA labs are more complete, consisting of items that were actually seized at scenes, during investigations, or from suspects or victims. In some places, forensic lab personnel routinely draw bloods from suspects, and a case scenario that creates an exception to this practice is required to construct a credible blind test. One of our blind proficiency test feasibility cases fell into this category, and the test was not detected. Agency and/or laboratory specific paperwork, as well as proper forms and evidence labels must usually accompany the evidence to the laboratory. In some jurisdictions, it is routine for the investigator

and the lab examiners to communicate about a case. Sometimes, the prosecutor's office must be included in discussions of the "case" and the submission early in the process (even to the degree of approving cases for DNA typing). In those instances, the prosecutors office must also be brought into the deception loop during the case setup.

(b) Types of Blind Proficiency Test Cases

In some ways, sexual assault "cases" are easier to manufacture than blood-transfer cases. Fewer items can often be submitted in a sexual assault case without arousing suspicion. On the other hand, more specimens may be needed from donors. Some sexual assault kits call for the collection of vaginal, oral and anal swabs, for example, and in some places, those specimens are collected even if there is no allegation by the complainant of oral or anal contact. Head and/or pubic hair specimens and/or combings may also be required. In blood transfer "cases" that would involve medium to high velocity blood spatter, the patterns on clothing and other items are often very difficult to simulate credibly under controlled laboratory conditions. It is also exceedingly difficult to manufacture truly replicate evidence specimens when complex blood-transfer patterns are involved.

(c) CODIS / Databasing Issues

Another set of issues that arises with blind proficiency test sexual assault "cases" is CODIS databasing of the profiles. First, there is the matter of initial donor informed consent and then of subsequent purging of the profile(s) after the test is completed. Second, there is the matter of case-to-case CODIS hits in the forensic files. Using the same male donor in tests for two laboratories that are both CODIS participants should yield a hit. In the short run, and on the small scale that our feasibility tests have been conducted, the CODIS issues have been controllable. Each cycle of the blind proficiency test feasibility tests contained a potential CODIS case-to-case forensic file "hit", and the hit was found in the pair of cases that got worked by the labs. One member of the other pair never got worked.

Looking ahead, however, any large-scale program of blind testing that intends to use a significant number of sexual assault "cases" will require a significant number of different semen donors if CODIS is up and running in most of the target labs and if the CODIS hit problem is to be avoided or controlled. Failure to control the number of case to case CODIS hits in multiple jurisdictions would likely compromise a whole cycle of blind testing.

7. Blind Proficiency Testing Feasibility - Summary

Four models for blind proficiency tests can be considered. Two of these, Blind / LE and Blind / CL, are fully "external," in the sense that no one in the target laboratory has to be involved in the design, manufacture or submission of the test "case." The Blind Analyst model could be "external" in the sense that the test case can be designed and manufactured externally, but people within the laboratory, other than the analysts, are involved at least in the broad outlines of the test design. These people might not know exactly when the test is submitted, or they might actually be involved in directing it to a particular analyst. Random Reanalysis is the fourth modality. In this model, a worked case is selected for reanalysis by another analyst or another laboratory, and this reanalysis is accompanied by a more or less complete "audit" of the targeted analyst's work in the case selected. Random Reanalysis could be "external" to the extent

that an outside "auditor" chose the case for reanalysis, and another laboratory reanalyzed the case evidence.

In Phase 1 of this project, we concentrated on attempting to show that Blind / LE and Blind / CL proficiency testing was possible in a number of different laboratories and jurisdictions. These fully external models for blind proficiency testing are the most complicated and challenging of the four. They are also the most costly. We have shown that fully external blind proficiency testing is possible, although we were not universally successful in our efforts—one of the "cases" was recognized as a blind, and another did not get worked. In the former case, we inadvertently ended up testing the Blind Analyst model, because the DNA analysts were not told that the case was a blind proficiency test by those who recognized that it was. Although this was only one case, it shows that the Blind Analyst model is also possible. Random Reanalysis, the fourth modality of blind proficiency testing, is obviously possible since a number of jurisdictions routinely use it as a part of their ongoing QA program.

The factors involved in planning and operating a proficiency testing program in general, and a blind proficiency testing program in particular, as well as the estimated costs of running such a program under the different possible models, are discussed in § V below.

B. Phase 2 Feasibility Trials

In Phase 2 of this project, additional blind trial proficiency tests were designed and submitted to forensic science laboratories. The objective of these Phase 2 tests was to gather preliminary data on the feasibility of the accurate replicate-manufacturing of case materials in cases that were more complicated than those used in Phase 1. The majority of biological evidence cases involve either blood transfer or sexual assault. These cases can be more complicated by factors such as blood mixtures, multiple bloodstains – that could be from different depositors—on evidence items, and semen mixtures on sexual assault evidence items. NIST has run a large, controlled mixture study involving accurately manufactured semen mixtures. In order to try not to duplicate that study, we decided to develop case scenarios and manufacture Phase 2 blind tests around a single evidence item on which had been deposited two persons' blood, that of the "victim" and that of the "suspect." In every case, the scenario involved assault, attempted sexual assault and/or home invasion, and sharp-force injuries (inflicted by a knife that was not recovered) to both parties. Some resulting bloodstains on the pants of the "victim" were from the "victim" while others were from the "suspect."

Our plan was to make the "cases" somewhat more challenging from a criminalistics point of view, that is, labs would have to make some decisions about which stains to examine. There was enough information about the case circumstances in every lab submission form to give an unmistakable signal that it might be necessary to type several bloodstains to locate those not deposited by the "victim." Five such blind "cases" were constructed and submitted to forensic DNA laboratories through law enforcement agencies.

Because part of the challenge in this phase was the reproducible, replicate manufacturing of evidence with bloodstains that had to exhibit a pattern consistent with the case scenario, we chose to manufacture a total of ten case items. Five were used in the blind test "cases" submitted to forensic labs, and five were submitted to reference laboratories. The reference laboratories received the same items as the test laboratories, but they were aware the cases were proficiency tests. Reference laboratories were given the same type of information about the "case" as was

given to the test-target labs in the lab submission forms, and they were told to examine the evidence as they would a regularly submitted case. Thus, we did not tell the reference labs in advance how many depositors there were, nor how many stains they should type.

All the labs participating in Phase 2 had agreed in advance to be potential test candidates (or reference laboratories), just as was the case in Phase 1. The Phase 2 tests were manufactured and administered in a single cycle.

Our survey of every laboratory that we thought might be performing forensic DNA analysis resulted in the return of 91 surveys (see § IIIB), 67 of which were actually doing DNA analysis. Of these, 32 agreed to be potential participants, but only 30 returned the signed "Agreement to Participate" forms. The potential participant pool thus consisted of 30 labs. These labs had the characteristics shown in Table IV-3.

Table IV-3. Characteristics of Target and Reference Laboratoeis – Phase 2

Number (%Total)	Type	Service Area	No. DNA Analysts (Avg)**	Cases Come Mainly From
16 (53.3)	State System	Entire State*	n = 17 (7.35)	Police / Other Labs
7 (23.3)	County	Entire County	n = 8 (2.75)	Police
4 (13.3)	City	Entire City	n = 5 (5.6)	Police
3 (10.0)	Private / Independent	No Limits	n = 6 (4.67)	Police, Attorneys, Other Labs

* not necessarily the sole provider of DNA typing services in the state
** n = number of labs who reported the number of DNA analysts

1. Selection of and Agreements with Participating Laboratories

Laboratories that agreed to be potential participants did so by signing a formal agreement with us (Appendix E1), containing certain mutual understandings. We gave them the same assurances as in the Phase 1 testing (see section § IVA above). And they in turn gave the same assurances to us as in Phase 1.

2. Blind Trial Proficiency Test Case Setup, Manufacture and Distribution

We subcontracted with the same TWGDAM-approved DNA proficiency test manufacturer used in Phase 1 to manufacture to our exact specifications the proficiency test evidence designed by the UIC project.

Approval was re-obtained from the UIC IRB for the employment of volunteer human subjects as donors of biological evidence specimens, and as to the substance and form of the informed consent statement. The informed consent statement was again provided to the manufacturing contractor for use with the volunteer donors (Appendix F). The specimens obtained from the donors were anonymized by the contractor, and the signed consent returned, sealed, to one of the UIC project directors. Knowledge of the identity of the volunteer donors was thus restricted in effect to one UIC project director.

Prior to the actual manufacturing of biological evidence, "cases" were set up with knowledgeable people in law enforcement agencies who would serve as the submitters. This step involved interviews to determine the detailed characteristics of a "typical" case of the kind we

had in mind for Phase 2, what evidence would typically be submitted, and in what form, how it would be packaged, how the packaging would be marked and by whom, what paperwork would accompany it, etc. All the "cases could be designed around the submission of three items: a pair of pants or sweat pants worn by the "victim", a "victim" exemplar and a "suspect" exemplar. If the jurisdiction made use of a particular collection kit, blood collection tube, etc., the evidence was submitted in the proper container. For most of the cases, a female "victim" and a male "suspect" were used. In one jurisdiction, it was necessary to have a male "victim." Thus, three specimen donors – two males and one female – were required for the Phase 2 case manufacturing.

Suspects and victims were assigned fictitious names, dates of birth, and sometimes race, and an offense date that made sense in terms of the normal submission lag time.

The detailed setup agreement with the submitter, and the case scenario, then enabled us to prepare detailed instructions for the manufacturer to put the biological evidence together. Complete specifications for a "case" included detailed instructions for the collection of specimens from donors, for use of those specimens in manufacturing the biological evidence, and for transmitting the items to the submitting entity. We also included a document to be sent to the submitting law enforcement agency or conduit lab along with the "case." In some instances, additional instructions for marking the evidence, etc., was included with this document. An example of these detailed instructions and specifications for Phase 1 is shown in Appendix G.

As in Phase 1, agreements were made and signed with law enforcement agencies that participated as submitters in this project (Appendix E-2 for a typical agreement).

As noted earlier, replicate bloodstained pants and exemplar specimens were sent to five reference laboratories for DNA typing about the same time the "case" specimens were forwarded to submitters. They were told the scenario, but were not told how many stains to examine or type. Because there were two separate case scenarios, one involving a "female" victim and the other involving a "male" victim, three reference labs received one of the scenarios and the other two received the second. The scheme is summarized in Table IV-4 below.

BT2-M1, -M2, and –F1 stand for "male 1," "male 2," and "female 1," respectively. The manufacturing contractor maintained detailed and meticulous records of the specimen collection, testing and use in the manufacture of blind proficiency test cases. This record keeping constituted a significant part of the contractor's QA procedures for manufacturing the specimens and cases. Samples of all the biological specimens used in case manufacturing were retained during the project period as well, in case any question about a specimen arose in the future.

Figure IV-1 (at the end of this section) shows representative images of the pants / sweat gear used for the deposition of bloodstains in this series of blinds. As with Phase 1, the target lab agency's normal forms were used for case submission for the Phase 2 blind tests.

Table IV-4. Scheme of the Phase 2 Blind Proficiency Tests

Blind Tests	BT2-M1 Suspect	BT2-M2 Victim	BT2-F1 Victim	Exemplar
11	minor	major	---	M1 and M2
12	minor	---	major	M1 and F1
13	minor	---	major	M1 and F1
14	minor	---	major	M1 and F1
15	minor	---	major	M1 and F1
Reference				
1	minor	major	---	all three
2	minor	major	---	all three
3	minor	---	major	all three
4	minor	---	major	all three
5	minor	---	major	all three

major: approximately 6.–7 drip stains + satellites (7 – 9 total)
smear right pocket, or "right pocket" area
minor: 1 drip stain + satellites (2 – 3 max)

3. Results

(a) Details of the Actual Blind Trial Feasibility Tests

As noted, five separate blind proficiency tests were carried out during Phase 2, and five replicate-manufactured "cases" were submitted to reference laboratories. All were blood transfer cases, and required three biological specimen donors.

In phase1, ten blind tests were set up, constructed and submitted (that number includes our collaboration with the contractor responsible for submitting blind DNA proficiency tests to the FBI Laboratory). Five volunteer donors contributed biological specimens to the first ten tests.

The characteristics of the ten Phase 1 blind proficiency tests were shown in Table IV-2 above. In Table IV-5 below are the characteristics of all fifteen blind tests from both phases of the project.

Table IV-5. Characteristics of Fifteen Blind Tests in Phases 1 and 2

Test	Target Lab Type	Submission through[a]	Type of Case[b]	Type DNA Testing Done[c]	Reported Findings	Turnaround Time (m)[d]
1	Private	CL	SA	PCR	Suspect included	0.83
2	Private	CL	SA	PCR, STR	Suspect included	1.63
3	State	CL	BT	RFLP	Blood on suspect clothing consistent with victim	5.06
4	State	LEA	SA	RFLP	Suspect included	6.9
5	Federal	LEA	SA	RFLP	Suspect included	16.53[e]
6	State	LEA	SA	PCR	Suspect included	3.46
7	Municipal	LEA	SA	RFLP	Suspect included	1.00
8	State	LEA	BT	N/A	Not completed[f]	3
9	State	LEA	SA	RFLP	Suspect included	4.06
10	Federal	CL	SA	RFLP	Suspect included	6.36
11	State	LEA	BT	PCR	Suspect included	4.5
12	State	LEA	BT	STR	Suspect included	11.4
13	Municipal	LEA	BT	STR	Suspect included	3.1
14	Municipal	LEA	BT	PCR	Suspect included	9.5
15	State	LEA	BT	PCR	Suspect included	2.1

Table Footnotes:
[a] CL: conduit laboratory; LEA: law enforcement agency
[b] SA: sexual assault; BT: blood transfer
[c] RFLP; PCR: HLA-DQA1, PM and sometimes D1S80; STR, various combinations of loci; N/A: not applicable
[d] Months, obtained by consistently dividing turnaround time in days by 30
[e] Hairs and fibers unit completed its report with a turnaround time of 2.16 m
[f] Lab request to police for additional specimen not communicated in a timely way to project team; see in § IVA.4(b) above

Phase 2 case manufacture was completed around mid-April, 1999, and the cases were all submitted within a few days.

(b) Revelation of a Blind Test in the Phase 2 Feasibility Trials

At one of the locations where a municipal laboratory was to be targeted, the project directors met with a police officer to set up the scenario and make arrangements for this police officer to submit the case once it was manufactured. The meeting itself was unremarkable, not unlike many others held for the same purpose, and took place at a location well removed from the agency building. A short time after the meeting, we received information from an otherwise uninvolved party that this police officer had revealed plans for the blind test to the laboratory director. The police agency also somehow became aware of this action, but not because we told them. A different person from the police agency contacted us, and indicated the department's willingness to go ahead and submit the case as originally planned. We then went ahead.

It turns out the blindness of the test was not compromised because the laboratory administration assumed we had changed our minds and submitted the case to another laboratory upon learning of the revelation.

We do not know why the police officer revealed the existence of the test plan after having agreed to handle the submission. There is no question the officer understood the necessity for confidentiality until after the test was submitted and completed. The experience did show that blind tests intended for submission through law enforcement can be compromised.

(c) DNA Typing Results

All the labs reported the correct results, in the sense of including the possible depositor of biological evidence, and all test-target labs detected the "suspect's" (minor contributor to the pattern) bloodstain. So did four of the five reference laboratories submitted comparable results. The project directors take responsibility for the one that did not. Communications with that laboratory during the set-up phase were with supervisory personnel who might or might not have conveyed the case scenario in the same manner we would have done, had we spoken directly to the analysts or immediate supervisors.

Three target labs did PCR-based loci HLA-DQA1, PM (and sometimes D1S80), and two did STR loci, reflecting the transition from HLA-DQA1/PM to STR loci taking place in the period when these tests were submitted and worked.

Four reference labs typed HLA-DQA1 and PM loci, one did D1S80, one did six RFLP loci. For the STR loci, two labs did eight loci and three did thirteen plus amelogenin. The reference lab data along with the target lab data for Phase 1 tests is collected together in Appendix H, and similar results for the Phase 2 testing can be found in Appendix H. Reference lab data was shared with every target lab following the conclusion of the blind tests.

(d) Turnaround Times and CODIS Issues

As can be seen in the table above, turnaround times in the Phase 2 tests varied from a little over 2 to a little over 11 months. The variation probably reflects the individual lab situations, including factors such as number of DNA cases in the jurisdiction, case backlogs, seriousness of the cases, etc.

Most of the target lab jurisdictions would not have databased the bloodstain results under their existing laws, and several as noted did not type databaseable loci. No CODIS hits were planned for this phase, and none occurred.

4. Phase 2 Summary

In Phase 1 of the project, we concentrated on conceptualizing blind-test models, and running small scale trials of the fully "external" ones (Blind/LE and Blind/CL). It was shown that such testing is possible, although detection of and problems with the tests do occur. In Phase 2, we focused attention on whether evidence items representative of more challenging cases could be replicate manufactured with sufficient reliability to insure uniform results from competent laboratories. This task was accomplished. At least insofar as relatively uncomplicated sets of bloodstain patterns on items are concerned, replicate evidence manufacturing is possible, although it is labor intensive.

We did not set out to make the Phase 2 cases overly complicated or difficult. We wanted to see what results would be obtained when judgment was required about the selection of evidence

for analysis and typing under conditions where the police description of the events suggested more than one blood source.

The results indicate that such blind tests can be constructed and successfully submitted to forensic DNA laboratories.

Figure IV-1. Images Showing Bloodstain Patterns on Manufactured Evidence

V FACTORS IN DEVELOPING A BLIND PROFICIENCY TESTING PROGRAM

A number of factors have to be considered in developing any proficiency testing program. Some of these are peculiar to a blind, as opposed to an open or declared, testing program. Both types of factors are discussed below to create the context for our discussion of the logistics and estimated costs of a large scale blind proficiency testing program.

A. Purpose of Proficiency Testing

Proficiency testing is ordinarily considered to be one component of a laboratory's QA program. Other components can include (i) examiner training and competency; (ii) ongoing continuing education and training of examiners as technologies change and evolve; (iii) use of validated methods and documentation of the validation; (iv) documentation of methods used in casework; (v) documentation of procedures used in any by notes / worksheets; (vi) routine procedures for and documentation of QC of chemicals, reagents and instruments used; (vii) routine use and documentation of internal controls in each step of each procedures; (viii) routine supervisory review of analysts' work, results and interpretation; (ix) routine supervisory review of analysts' expert testimony; and (x) full laboratory audits by outside experts that typically include review of all or many of the foregoing factors.

Proficiency testing can test a number of different components of individual analyst and/or laboratory performance in working a biological-evidence (or any other) case from start to finish. These components include: (i) analytical results; (ii) interpretation of analytical results; (iii) all records, worksheets, and documentation; (iv) compliance of performance in the test case with the laboratory's established procedures; (v) compliance of the report with laboratory policy; and (vi) evidence selection, sampling, and other judgments made in choosing specimens for DNA analysis in the context of the "case." Different proficiency-testing modalities have strengths or weaknesses with respect to how well they provide a basis for measuring performance in these various components.

B. Declared vs Blind Proficiency Testing

A full-scale, obligatory declared proficiency testing program is already in place in all DNA testing laboratories. That is, forensic laboratories may choose from several vendors that supply declared DNA proficiency tests. There are many components of a QA program besides proficiency testing, in addition to the various QC procedures laboratories normally practice as part of their routine casework, as noted above. There also appears to be an underlying assumption that the purpose of proficiency testing is to see how examiners are performing as compared with most examiners in most laboratories. Accordingly, current proficiency testing programs tend to issue tests that are typically designed with the expectation that the majority of participants will produce an acceptable result or response. We refer to this kind of test design as "testing to the average."

C. The Case For and Against Blind Proficiency Testing

The case for blind (instead of, or in addition to declared proficiency testing) is based on the information that might be obtained from blind proficiency tests that is not available from declared proficiency tests. Two points are generally cited in support of blind vs declared proficiency testing.

First, there is evidence from the clinical proficiency testing literature that examiners on the whole will perform better on declared proficiency tests simply because they know they are being tested. In regulated clinical laboratories, open proficiency testing is a major, if not the sole, criterion for a laboratory retaining its license to do particular tests. Here, analysts must also attest to the fact that they used their 'standard' procedures in examining the sample and that they did not collaborate with other analysts/laboratories. As such, proficiency tests are used more as regulatory tools in this environment than as educational devices designed primarily to improve quality. It is unclear to what extent this fact is responsible for the better performance in declared vs blind tests. In the context of forensic DNA laboratories, it must be kept in mind that while declared proficiency tests can be directed to individual examiners (to the extent that laboratory division of labor permits it), blind proficiency tests submitted through LE agencies or CLs cannot. The Blind Analyst and Random Reanalysis models do permit particular examiners to be singled out for testing to the extent possible in an individual laboratory.

Second, it is often said that blind proficiency testing tests "the whole system" whereas declared proficiency testing primarily tests the ability of the participant to obtain an acceptable analytical result, and maybe an acceptable interpretation of the results. By "the whole system" in this context is meant all the steps and record keeping that go into case intake, sorting and selection of items for analysis, screening or preliminary tests, DNA analysis itself, interpretation of the results and preparation of a report. To the extent, therefore, that it is desirable to test all these aspects of the forensic lab analysis "system," in addition to the acceptability of the analytical results, as part of an ongoing QA program, blind proficiency testing would be required with some suitable frequency.

Compared with declared testing, blind proficiency testing is complicated and expensive under the fully external Blind / LE model, arguably the only one of the blind models that completely fulfills the criterion of testing "the whole system." All the blind testing models satisfy the criterion of testing examiners without their knowledge.

Laboratory administrators, QA administrators and policy-making bodies who have the authority to mandate forensic DNA lab QA procedures need to decide whether the additional information obtained from a blind proficiency testing program, as compared with the existing declared program, justifies the additional effort and costs associated with introducing a blind program. There will doubtless be disagreements about this issue. To some extent, those disagreements result from differing viewpoints about the goals and purposes of proficiency testing as articulated above.

D. Defining Acceptable Performance / Performance Review

In any proficiency testing program, two ingredients are essential if the program is to serve its avowed purpose as a meaningful QA component. First, there must be some way of defining an acceptable response or result; and second, there must be a review of the laboratory's and/or analyst's performance in the test

1. Acceptable Performance

In any proficiency testing program, participating laboratories have to know the criteria that will be employed in defining acceptable performance. Some of the performance components enumerated above, such as compliance of various aspects of test procedures and interpretation with the laboratory's established policies, are relatively easy to judge. Others are less so.

In terms of obtaining acceptable analytical results, there are generally three criteria that can be used in proficiency testing: (i) results consistent with manufacturers specifications; (ii) results in concordance with one or more reference laboratory results; or (iii) results in accordance with a consensus value or response based on the pool of tested laboratories. The first of these is not applicable to tests involving a laboratory's ability to type genetic characteristics. In the context of DNA typing, acceptable analytical results could be defined using either of the second two criteria, but the definitions will differ depending on whether RFLP typing or discrete allele locus typing has been done. With discrete genetic system typing, an acceptable result is normally one that agrees with one or more reference laboratory results, or with a consensus value. With RFLP typing, an acceptable result is normally one that lies within either (i) a predefined range based on the mean values produced by reference laboratories, and calculated as either a percentage of mean values or some fraction or multiple of standard deviation; or (ii) a predefined range based on the mean values produced by consensus of all tested labs in that cycle, and calculated as either a percentage of mean value or as some fraction or multiple of standard deviation.

Other factors in addition to the analytical result itself might be used to judge whether a result is "acceptable." Among these could be: (i) Was sufficient DNA obtained from the specimen(s), given that some certain amount was there; (ii) Were the frequencies of the DNA profiles properly computed and cited where applicable; (iii) If a comparison was called for, was the interpretation of the types in terms on including or excluding possible depositors satisfactorily rendered -- this factor could be particularly important in mixture cases; or (iv) Were appropriate judgments made in the context of the case in terms of what evidence items should be typed. With these factors, it may be more difficult to set the criteria for an acceptable response from the tested laboratories. Suppose, for example, a lab obtained too little DNA for RFLP typing when most of the participating labs or the reference labs did so. Using this factor as a measure of lab performance gets into issues of extraction efficiency, the lab's human DNA quantitation procedures and their efficiency, and the criteria a lab uses to judge whether it has sufficient DNA for RFLP typing (that is, its lower limits of RFLP test sensitivity). In addition, the ability of the test manufacturer to prepare truly replicate test specimens are an issue in this context – a lab that did not obtain as much DNA as others might argue that the particular specimen they received actually had less to begin with.

In general, but especially if laboratories will be subject to sanctions for "unacceptable" performance, it is very important that the criteria for acceptable performance in the proficiency testing program be made plain at the outset. If the purpose of proficiency testing is to see how examiners are performing as compared with most examiners in most laboratories, there must be some predetermined consensus by those who will judge the results of a proficiency test as to what will be expected. If factors other than the analytical results themselves are to be considered, there must be agreement on what those factors will be. Further, program participants should be made aware of what the factors are, and how the test performance evaluators will use them in defining an acceptable result or response. The extent to which factors other than analytical results themselves will be considered in defining an acceptable result or response will, in turn, influence the nature and complexity of the proficiency testing specimens and requirements.

2. Performance Review

Laboratory administration typically reviews proficiency test performance by its examiners, and there are written guidelines within laboratories for the handling of less than satisfactory responses. Errors are inevitable in any type of laboratory testing and guidelines must be in place that define the types and/or frequency of different types of mistakes that will trigger some type of remedial/corrective action. In some situations, an isolated error may evoke a response where with others the review may seek to identify a pattern of unacceptable responses. Sanctions may focus on individual analysts , sections, or entire laboratories. Proficiency testing error handling is further discussed in §V.E below.

ASCLD-LAB's Proficiency Test Review Committee looks at the performance of accredited labs. ABC-certified criminalists are required to report the results of their proficiency tests for re-certification.

For a large scale program, a good argument can be made for having the same group that set the criteria for test performance be the entity that evaluates performance in the tests for purposes of setting national standards and guidelines (see in V.9.a below).

E. Errors / Error Handling

It is essential, as noted above, that the factors proficiency-testing judges intend to use in defining acceptable results / responses to proficiency tests be specified in advance. Further, program participants need to know whether there are predefined levels of acceptability in any of the factors that will be used in judging test performance. The proficiency-testing reviewers might decide, for example, that a turnaround time of greater than 8 weeks for a case was unacceptable. Thus, a laboratory's results / response might be judged unacceptable because it was too slow, even though every other aspect of it was perfect. Not only do participants have to be made aware of factors of this kind, the public documentation of the proficiency testing program should clearly explain the different bases and criteria used to define an acceptable result. No one would seriously argue that an "unacceptable" rating in a proficiency test because of being two weeks late in reporting is in any way equivalent to an "unacceptable" rating in a proficiency test because of a demonstrable analytical or interpretation error. If it is decided that multiple factors are to be used in judging acceptable results or responses, it would be essential to recognize and distinguish qualitatively different types of "errors" or "unacceptable" responses, and communicate the differences to consumers of test results.

Another important aspect of error definition is how errors or unacceptable responses will be handled by laboratories participating in the testing program. Just as the QA guidelines promulgated by TWGDAM, DAB, etc. call for labs to have written, consistent procedures by which errors in declared proficiency tests are handled, the same or modified procedures must be in place for blind proficiency tests, should such a program be a voluntary or mandated part of the laboratory's QA program. The procedures for responding to errors or unacceptable responses in blind proficiency tests may have to be more extensive than the parallel procedures for declared proficiency tests, because a wider range of lab procedures are "tested" in a blind.

A final point concerning errors has to do with the question of error rates. The issue of error rates in analytical testing was raised in the Supreme Court's 1993 Daubert decision on evidence admissibility. To date, the frequency of proficiency testing has been so low in comparison with the frequency of casework specimen handling, that error rates in proficiency tests cannot be considered representative of casework practice and, therefore, cannot be directly extrapolated to

represent overall laboratory error rates. Representativeness also requires that the proficiency tests mirror actual cases (in terms of their realism), and that the proficiency test environment reasonably approximates the conditions under which actual cases are examined. Presumably, blind tests are more representative of actual cases than are declared tests, both in terms of their analytical requirements as well as of their forensic / criminalistics features. An added requirement for error rates to be meaningful at the national or profession-wide level would be the assurance that the proficiency tests themselves are of comparable difficulty and that the testing conditions within the laboratories, including whether the tests are declared or blind, are comparable. Otherwise, the merging of proficiency test data would not be meaningful. At this stage, though, it is highly unlikely that declared or blind proficiency tests can be administered with sufficient frequency, relative to the quantity of most laboratories' caseloads, to allow a meaningful calculation of the lab's or any examiner's "error rate."

F. Blind Proficiency Testing: Introduction Modalities and Internal vs External

Subtitle C of the DNA Identification Act of 1994 (P.L.103-322; Appendix A) referred to blind external proficiency testing, and defined it as "... a test presented to a forensic laboratory through a second agency and appears to the analyst to involve routine evidence." In this discussion of blind proficiency testing, therefore, the matter of internal vs external must be discussed, and ultimately defined, along with the four different modalities for the introduction of blind tests to a laboratory.

As discussed in §IV.6(a) above, there are four distinguishable modes of introducing or administering blind proficiency tests to forensic DNA laboratories. We have called these: Blind / LE; Blind / CL; Blind Analyst; and Random Reanalysis. They are fully defined below.

1. Introduction Modalities

(a) Blind / LE

[Blind to everyone in the target lab; submitted via law enforcement agency] No one associated with the target laboratory is informed/aware the test is to take place. The only contact is made with an external, law enforcement agency that collaborates with the testing organization to create the specimen, the necessary case report, and the related paper work. This is the most challenging of all models for the testing organization.

(b) Blind / CL

[Blind to everyone in the target lab; submitted via another lab that does not do DNA typing itself -- conduit lab] No one in the target laboratory is informed/aware of the test, but another laboratory is involved in submitting the blind proficiency test case items to the target lab. The conduit lab may be part of the same lab system as the target, or completely independent of it (e.g., if a public forensic laboratory submits items to a private or independent DNA typing lab). There is no difference in the level of blindness, whether the conduit lab is or is not part of the same lab system. In terms of level of blindness, there is no difference between the Blind / LE and the Blind / CL models.

(c) Blind Analyst

[Only the analyst is blinded] In this model, the bench analyst is unaware of the test, but someone in the laboratory (director, supervisor and/or QA coordinator) knows, and may be involved in the construction of the test/case. It might be argued that there is greater potential for the "blindness" of a test using this model to be compromised. We did not design any Blind Analyst model tests in the present studies, but one Blind / LE test, though detected in initial screening, was sent on for DNA analysis anyway. So one test was inadvertently conducted in this way, and the DNA analysts were kept blinded until the test was completed. Critics of the labs might find more grounds for challenging the integrity of the "blindness" of a test under this model than under one of the foregoing ones.

(d) Random Reanalysis

Case audit with reanalysis of previously tested case material is a model that can achieve many of the same purposes as other types of blind proficiency testing. Here, the biological evidence from a completed case is re-tested by another laboratory. In addition, all the paperwork, worksheets and other work products from the case is typically reviewed and critiqued. Several laboratories use this model as part of their ongoing QA/QC program. An analyst (or several analysts) is (are) tested blind in this model, because they do not know which case may be selected for audit/reanalysis.

The random reanalysis model is used in larger laboratories and in laboratory systems where more resources are available to devote to it. In one version of random reanalysis, a designated examiner, acting as a quality coordinator, assists the QA manager in selecting cases for audit/reanalysis. The constraints are that the case must have been worked but still be in the laboratory, and there must be sufficient remaining specimen to permit reanalysis without consuming everything. Upon completion of the audit/reanalysis, the user agency is notified of the results if a report has already been issued in the case. If a problem is discovered, it is taken up with the original analyst, QA manager, and laboratory director. If the problem is sufficiently serious that its correction alters the conclusions or interpretations, an amended report is issued if necessary.

From a national perspective, dependence on a random reanalysis program is problematic if national inter-laboratory comparisons are desirable, in that it would be extremely difficult to insure comparability among the cases reexamined. The composition and level of difficulty of samples could not be standardized a priori like manufactured proficiency tests, nor could the presence of possible contaminants be controlled. Consequently, the pooling of results to construct a profession-wide profile might not be possible. It should also be appreciated that the exact same specimen analyzed initially cannot be reanalyzed since it was, by definition, consumed. This factor will not usually be a problem, but it could be if a case involved complex stains consisting of partially overlaid mixtures.

2. Internal vs External

The distinction between external and internal blind proficiency testing in terms of the four different blind testing models is not as simple as it might first appear. All four introduction modalities could be conducted (and thus defined) as internal or external, depending on the details of the way a test is conducted (see Table V-1 below).

Table V-1. Blind Test Introduction Modality and Factors Defining the Test as Internal vs External

Blind Test Introduction Modality	Factors Tending to Define the Test as External	Factors Tending to Define the Test as Internal
Blind / LE	• No one in the lab has any advance knowledge of the test • Test is manufactured externally and some external testing agency arranges for submission of the test through a LE agency that regularly submits cases to the lab	• Director, supervisor or QA coordinator orders the test, and may advise the external testing agency on type of case and details • The individuals "in-the-know" might know which incoming case is a test • Test could be manufactured internally and someone arranges for submission of the test through a LE agency that regularly submits cases to the lab
Blind / CL	• No one in the lab has any advance knowledge of the test • Test is manufactured externally and some external testing agency arranges for submission of the test through a LE agency that regularly submits cases to the lab	• Director, supervisor or QA coordinator orders the test, and may advise the external testing agency and/or CL on type of case and details • The individuals "in-the-know" might know which incoming case is a test • Test could be manufactured internally and someone arranges for submission of the test through a CL that regularly submits cases to the lab
Blind Analyst	• Director, supervisor or QA coordinator orders the test, and may advise the test preparer on type of case and details • Test is manufactured externally and some external entity provides the case for submission to the DNA analysts	• Director, supervisor or QA coordinator orders the test • Test might be prepared internally • The individuals "in-the-know" arrange for submission of the test to the DNA analysts
Random Reanalysis	• Director, supervisor or QA coordinator orders the test • An external auditor selects the case for reanalysis • An external auditor reviews every aspect of the case from start to finish, including a review of the selection of evidence items for analysis • Evidence items are reanalyzed for DNA types by an external laboratory-other laboratories part of the same system could be excluded	• Director, supervisor or QA coordinator orders the test • Director, supervisor or QA coordinator selects the case for reanalysis • Director, supervisor or QA coordinator reviews every aspect of the case from start to finish, including a review of the selection of evidence items for analysis • Evidence items are reanalyzed for DNA types by a different analyst or another laboratory in the same system

Thus, any of the four introduction modalities for blind proficiency tests could be considered either external or internal, depending on the details of how the testing was arranged and conducted, particularly in terms of what we have called "levels of blindness." Accordingly, a definition for "external blind proficiency testing" is necessarily somewhat arbitrary.

For purposes of this project and report, we define an external blind proficiency test as: (1) A test presented to a target lab through law enforcement or a conduit lab in which the "case" or "evidence" was externally manufactured, and no one in the target lab has any advance information about the test; or (2) A test presented to the DNA analysis unit in which the "case" or "evidence" was externally manufactured, and in which the fewest possible personnel outside the DNA unit are informed about the test; or (3) A test by "random reanalysis" in which auditors / analysts from outside the laboratory (and outside the laboratory system if the lab is part of a system) select the case for reanalysis, audit / review all the work done in the case, and reanalyze the biological evidence.

If we assume that the value of blind tests, versus declared tests, is worth the added expense in terms of their realism, we must also consider the need/desirability of external blind tests. It is generally stated that the desirability of external blind proficiency tests over internal ones lies in the fact that laboratory administrators' involvement somehow compromises the integrity of the tests. That is, it assumes that there is a greater likelihood that the knowledge of the case being a test will remain protected with the involvement of a law enforcement agency in the loop, than if the case were to be administered by the laboratory itself. In other words, it assumes a police officer or investigator is less likely to tip off a laboratory or DNA analyst that a case is actually a proficiency test than is a laboratory administrator or QA supervisor. Although such an assumption may not be warranted – there is no evidence for it -- one might also argue that a blind external test at least gives the appearance of greater integrity than a test administered by the laboratory itself. One should not forget that the clinical laboratory field opted not to follow such a path and, instead, administers declared tests, and relies on the integrity of the individual examiner to verify that routine procedures were followed.

G. Characteristics of Blind Proficiency Tests According to Introduction Modality

As noted in §IV, the different modalities of introducing blind proficiency tests into DNA laboratories represent different "levels" of blindness for the target laboratory. Further, as noted in the foregoing section (§V.F), any of these models can be defined as "internal" or as "external," depending on the way the testing in the model is implemented. Here we discuss the way in which these test introduction / administration modalities involve different logistics and manufacturing, complexity that ultimately affect costs.

Blind / LE are tests where no one in the target laboratory is informed/aware the test is to take place. The only contact by test administrators is made with an external, law enforcement agency that collaborates with the testing organization to create the case scenario, specimens, the necessary case report, and the related paper work. This is the most challenging of all the modalities for the testing organization. Blind / CL represents the same level of blindness to the target laboratory, but the blind proficiency testing case is submitted through a conduit lab. The conduit lab might or might not be part of the same lab system as the target. Blind / CL proficiency tests are less difficult to manufacture and introduce than Blind / LE proficiency tests. In the Blind Analyst modality, one or more people in the laboratory (director, supervisors, etc.) know about the proficiency test, but it is blind to the DNA analyst(s). Depending on how a laboratory operates, and especially how specialized the analysts are, this introduction modality

could involve different types of blind proficiency test case preparation in different labs, but this is probably the easiest and least costly of the externally supplied / introduced blind proficiency testing possibilities. Random Reanalysis is simply a reanalysis and "audit" of a previously tested case, where there is adequate evidentiary material remaining to enable the analytical steps to be repeated. Random reanalysis can have all the features of the other modalities of blind testing, can be directed at a particular analyst or section -- depending on how the DNA unit is set up, and could be the least costly of the blind proficiency testing models in terms of out-of-pocket laboratory expenses (see §V.J below).

In the Table V-2, the information items obtainable from proficiency testing, some of which are generally cited as justification for blind, as against declared, proficiency testing, are shown along with the extent to which each of the proficiency testing models supports the information item.

Table V-2. Information Items Obtainable from Proficiency Testing

Proficiency Test Information Item Or Property	Declared Proficiency Tests	Blind Proficiency Test Models			
		Blind / LE	Blind / CL	Blind Analyst	Random Reanalysis
Tests analytical results	+	+	+	+	+
Tests interpretation of analytical results	+	+	+	+	+
Tests (audits) all records, worksheets, paperwork	-	+	+	+/- [a]	+
Tests compliance of procedures with lab policy	+/- [b]	+	+	+	+
Tests compliance of lab report with lab policy	+/-	+	+	+	+
Tests evidence selection, sampling, and other judgments made in choosing specimens for DNA analysis	-	+	-	+/- [c]	+/- [d]
Allows construction of very challenging tests (Tests at the margin) [e]	+	+	+	+	+/- [f]
Likelihood of detection by target laboratory	N/A	Higher	Lower	Lower	N/A
Cost [e,g]	+	++++	+++	+++	++

Table Footnotes:

\+ the particular model supports the information item

\- the particular model does not support the information item

+/- the particular model may or may not support the information item

[a] Test documentation and record keeping are tested; evidence intake records and documentation are not

[b] Lab could follow the same procedures for the test that are followed for casework; procedures could then be reviewed

[c] Only if the DNA analyst were responsible for handling the evidence, and the "test" was made to look like a case

[d] This information item could be "tested" by audit of the procedure followed, provided the original evidence was still available for inspection by the auditor or reviewer

[e] Discussed below

[f] Although the "test" is not constructed, complex or challenging cases could be selected for random reanalysis

[g] Number of + signs indicates a crude measure of relative cost.

Table V-2 provides a summary representation of the information that can be obtained by proficiency testing, whether open (declared) or blind using the different models we have discussed. The two-part key question for laboratory QA administrators and for policy-making entities that have the power to impose QA procedures on laboratories is: (1) Is enough important additional information obtained from blind testing, as against open testing, to justify the additional costs? And (2) If a decision is made to implement blind proficiency testing, should a particular testing model be recommended or required?

Generally, two reasons are cited for doing blind, as against open, proficiency tests. First, labs may perform better in open testing, because they know they are being tested. Another way of saying the same thing is: blind testing results will more closely resemble performance in actual casework. There is some evidence from the clinical proficiency testing literature that labs do perform better in open proficiency tests. What is not clear is whether the regulated clinical laboratory environment is comparable to the forensic laboratory environment. Second, blind proficiency testing can test "the whole system," rather than just the analytical results. In the foregoing table, we tried to articulate some of the features of "the whole system" that are amenable to testing, and then compare the ability of various proficiency testing modalities to test them.

Nothing in our feasibility studies, nor in our review of the forensic proficiency testing literature, gives a strong indication one way or the other as to whether laboratories will perform better in open tests than they do in blind tests. The evidence that comes closest to addressing this issue directly is from the forensic urine drug testing proficiency test literature. Labs tended to find drugs present at sub-threshold levels in open tests, whereas in blind tests they did not detect those same drugs. The apparent reason is that the labs either took more time to look for drugs present in small quantities in the specimens they knew were proficiency tests than they did in specimens that they thought were routine, or they simply did not report the presence of drugs in the blind tests that were present in sub-reporting-threshold quantities. Arguably, however, it doesn't really matter, since the drugs detected at sub-threshold levels in the open tests did not meet the reporting threshold, and would not have been reported in case specimens even if they were detected.

As to "testing the whole system," random reanalysis can be the least expensive blind proficiency testing modality in terms of laboratory out-of-pocket cost. And this modality can address all the information items obtainable from blind testing.

One of the features that some modalities of blind testing can measure, and that cannot be readily be measured by open testing, is the initial judgments made concerning the selection of evidence for genetic typing, and the interpretation of results in terms of the case, especially when the results are complicated -- this feature might be called the "criminalistics" of the case. Many DNA analysts are not trained as criminalists, and in many laboratories the criminalistics of a case is done by personnel other than DNA analysts. Blind testing designed to try to test the criminalistics of a case would thus have to be considered tests of the laboratory in most instances, rather than tests of the DNA unit or the DNA analysts.

The only way to "test" the criminalistics of a case by blind proficiency testing is to construct (or select, in the case of random reanalysis) cases that force analysts to try to locate and identify all the important evidence and/or make decisions about what evidence should be sent on for genetic typing. Another testing technique could be constructing tests that have biological evidence mixtures. If the tests were constructed, the manufacturer would presumably have

control over the relative levels of mixture components in the evidence stains. In this report, we refer to this type of proficiency testing as "testing to the margin." This sort of test is designed to be difficult or tricky in some analytical or interpretational respect, such that a significant number of respondents may give unacceptable results / responses. Proficiency tests are generally constructed "to the average" (§V.B above), i.e., they are designed such that most of the laboratories or analysts will get an acceptable response most of the time.

Testing to the margin can be a useful educational tool for laboratories. However, it is difficult to imagine how such testing could be used to test laboratory or analyst "proficiency." Tests at the margin are difficult in several respects. First, it would be very difficult to get consensus on what an "acceptable" result / response should be. If this consensus cannot be reached, the test program is of little value. It might be difficult even to get agreement on reference labs -- by definition, reference labs should obtain an "acceptable" result or response. Second, these blind proficiency tests will be very difficult to manufacture, especially in reliable replicates, and replicates are always necessary. Third, even if the foregoing obstacles can be overcome, many people familiar with proficiency testing have suggested that tests at the margin are not useful tools in assessing a laboratory's casework QA program. They can sometimes be useful tools in getting at the source of a widespread problem, one that might be shared by many laboratories, whether the problem is analytical or interpretational. That type of use of testing at the margin, however, is primarily educational, and is designed specifically to improve QA, lab performance, a particular analytical technique, or to help reach consensus on an interpretational issue. Almost by definition, a significant number of laboratories would not give "acceptable" results / responses to tests at the margin, either because there was considerable variability in the results and responses, or because an "acceptable" result / response could not be agreed upon. Thus, while this sort of testing can serve a useful purpose under some circumstances, there is not much doubt that in the U.S. adversarial legal system, the results of tests of this kind would be used as devices to criticize and discredit the labs in a way that would be very unfair. Tests at the margin do not represent routine casework evidence and problems.

A final point in this section should be made about the chances of a lab detecting a blind proficiency test. One of our feasibility trial tests, submitted through law enforcement, was detected by the laboratory at the case-screening stage. Thus, even with considerable care and attention to detail, blind tests in a large-scale program will probably be detected periodically. Intuitively, one would think that there is a greater chance of detecting tests in the Blind / LE model than in the Blind / CL or Blind Analyst models, because they are necessarily more complicated to manufacture, and have to have the look and feel of real cases. In the latter two models, evidence items rather than "intact cases" can normally be submitted.

H. The Home Office System Experience

The British Home Office laboratory system has conducted blind proficiency tests in all its laboratories for some years. A special unit in the system has the responsibility of preparing the tests, and the QA coordinators for the various laboratories and/or disciplines decide how many such tests -- and in what specialty areas -- will be administered each year. The program is not limited to biological evidence or DNA. Blind "cases" can involve many types of evidence, just as in real-world casework. The Home Office laboratory administration considers the program to be successful, and an important component of their overall QA system.

The "cases" are prepared and manufactured to exacting specifications, and submitted to the target laboratories through normal law enforcement channels. Over time, law enforcement units

have come to cooperate fully in this program. The task is undoubtedly made easier, in comparison with the U.S., by the fact that only about eight target labs are potentially involved, and that the British police are under a unified overall command.

Acceptable responses are established by the system's QA coordinators, and the laboratories know the criteria by which test performance and results will be judged.

Blind cases are constructed to resemble routine casework, and to make inter-laboratory performance comparisons possible, especially over several years time. Blind cases are rarely constructed to "test at the margin." When they are, there is generally a specific issue to be resolved or question to be answered, and the purpose of this testing is to see whether open discussion concerning the question or issue among the laboratories has satisfactorily resolved it.

I. Logistics of a Large-Scale Blind Proficiency Testing Program

1. Proficiency Test Review Mechanisms and Test Coordination

A question that needs to be decided in connection with a national blind proficiency testing program is whether there will be any coordinated, national level review of the results. If such review is deemed desirable or necessary, there are a couple of implications for a large-scale program. First, there is an impact on the overall costs of the program. Second, national level, periodic review of results is the only way that the program would serve the purpose of giving a national picture of performance levels.

Another issue that needs to be decided in connection with the implementation of a national blind proficiency testing program is whether to have some national coordinating entity for the tests. This entity could be, but does not have to be, the same group who might review the results. A coordinating entity is probably essential if the Blind / LE, Blind / CL or Blind Analyst proficiency test models are to be followed. Coordination of these blind tests would involve the planning of the tests, setting up appropriate contacts, setting out the manufacturing specifications for manufactured blind tests, deciding on the number and types of tests per lab per year, and setting the guidelines for acceptable responses. There are several ways such a national coordinating entity could function. It might designate, or set guidelines for the approval of, one or several blind test administrative organizations or groups, and provide policy guidance. In terms of blind test manufacturers, the national group could do something similar. The blind test manufacturers and administrators could be the same organization or group, but if they were not, coordination between them would be required. A blind test administrative entity should have a detailed knowledge of the operations of forensic science labs, case submission, evidence handling, and adequate knowledge of biological evidence analysis to be able to plan and execute credible blind tests. There are currently guidelines governing TWGDAM approval of proficiency test manufacturers. These or similar guidelines, perhaps set by the DAB, would be necessary to designate or approve blind proficiency test manufacturers.

There are various ways in which the test coordination / manufacturing functions could be implemented, and they differ mainly in the number of coordinating or manufacturing entities who are involved, and in the way they interact with one another. At one extreme, there could be a single blind test administrative unit that included test- manufacturing capability to serve all the laboratories. Toward the other end of the spectrum, there could be a number of blind test administrative units serving different groups of labs in a region, and several approved manufacturing entities, all potentially interacting with the different regional administrators. There are some advantages in having more centralized test coordination / administration and

manufacturing, but the extent to which these are real advantages goes back to what policy makers want to get out of the overall program. Table V-3 summarizes some of the advantages and disadvantages of more centralized vs less centralized coordination / administration / manufacture of tests in terms of some of the features of a large-scale testing program.

Table V-3. Advantages and Disadvantages of More / Less Centralized Proficiency Test Coordination / Administration / Manufacture

More Centralized	Feature or Property	Less Centralized
+	QA / QC of Test Designs	-
+	QA / QC of Test / Evidence Manufacturing	-
+/-	Costs of Testing [a]	+/-
+	Approval of Test Manufacturers	-
+	Standardization of Proficiency Tests / Testing	-
+/-	Control Over Number of Donors / CODIS Issues [b]	+/-

Table footnotes:

+ in the column means that there is an advantage to the model in terms of the feature of property; - means there is a disadvantage. +/- indicates that a broad stroke conclusion about whether the model is an advantage or disadvantage is not possible without knowing the full details of the model to be followed

[a] Depends on the exact model followed

[b] Centralized model provides the most control; but could be managed under a regional model

Almost all the considerations discussed just above are applicable, as noted, to Blind / LE, Blind / CL or Blind Analyst testing models. National coordination or oversight of a blind proficiency test program relying solely on Random Reanalysis is much simpler. There is no manufacturing involved, and there is no need for any test-administration entities.

2. Logistics According to Blind Test Modality

The logistics of operating a program that would include all DNA testing laboratories in blind proficiency testing following the Blind / LE or Blind / CL (externally submitted and lab fully blind) models even once annually are intimidating. These testing models require a central test administrator, or several test administrators whose operations are overseen by a policy setting group. It is likely that the test administrator would have to visit every LE agency or conduit lab that would be submitting the blind proficiency testing cases. Detailed discussions around what is actually required, the jurisdiction's or agency's usual practices, the required paperwork, etc. are necessary to plan credible case scenarios and manufacturing specifications. Discussions might also be required with sexual assault evidence collection personnel and with prosecutors' offices in some jurisdictions. An additional problem is turnaround time. In most places, blind proficiency test "cases" must be less significant so that analysts will not detect them because they weren't in the news. In many labs, the less significant cases take on a lower priority because of resource constraints. In our limited feasibility study, only two of fourteen completed

cases were done within a month, seven took from 1 to 6 months to be completed, and five took from 6 to over 16 months (see §IVB.3.a).

Voluntary cooperation from LE agencies or CLs is not a sure thing. Unless blind proficiency testing were required of the lab to meet a QA guideline or accreditation standard, and the lab could, in turn, convince law enforcement to cooperate, there would be many labs for which a suitable, cooperative submitting entity could not be identified. Another potential problem is clandestine revelation of the tests to the lab by law enforcement. In one of our feasibility tests, a police officer nominally cooperating with us revealed our plans to the laboratory (§IVB.3.b). In that case, we found out about the incident, but it could have been otherwise.

There will probably be about 150 DNA testing labs in the country when all the programs in developmental or planning stages are fully operational. Even if only one test were required annually, a test administrative entity would have to make about 150 site visits in 12 months (about 12 per month). We believe that these visits would be essential, at least for the first couple of years of the program, if the program were to be mostly successful in terms of externally introducing blind proficiency tests without detection.

We also believe from our limited experience in this project that a maximum of ten blind proficiency test cases can be manufactured at one time and still maintain rigorous QA in the process. Under the 150 lab once per year scenario, 15 cycles of case manufacturing would be required.

To the extent that sexual assault cases were used as bases for the blind tests, a large number of male donors would be required to avoid, or even control, the CODIS cross jurisdiction case to case hit rate, once the NDIS component of CODIS is fully functional. Until or unless cross-jurisdictional case to case CODIS hits become a common occurrence, multiple hits of this kind within a short period would immediately signal the blind proficiency testing cases. It is not clear that such a large pool of male donors is available, given that the downside of having their DNA profiles in databases for significant periods of time has to be thoroughly explained to the potential subjects. It would probably be necessary to recruit significant numbers of female donors as well. Although there is no databasing problem associated with their DNA profiles, labs could very well notice the periodic recurrence of a "victim" DNA profile in casework -- and this could be a tipoff that the case was a test.

As noted above, there are different ways the administration of a large-scale program based on Blind / LE or Blind / CL models could be handled. Here, greater centralization creates more efficiency and fewer problems with coordination. The more decentralized the test administration and manufacturing functions become, the more complicated the coordination of all the testing on a national scale will become.

Some logistical compromises are possible under the Blind Analyst model of blind proficiency testing. Here, only the DNA analyst(s) need be blind. Accordingly, lab administration or lab QA administration would be involved in planning blind proficiency tests. Depending on how the lab operated, especially with respect to how close to or far removed from the actual evidence the DNA analysts are, some Blind Analyst proficiency tests could be made up fairly easily. The further removed from evidence receipt, screening, preliminary testing, etc., the DNA analysts are, the easier it would be to introduce blind proficiency test evidence. The costs of Blind Analyst model proficiency testing would be significantly lower than with either of the external, fully blind modalities, because every lab would, in effect, have control over the

planning and implementation of testing of its own analysts. Generally, the biological evidence items could be simpler and less elaborate, since in many cases analysts would be expecting cuttings, bloodstain swatches, etc. In the Blind Analyst model, blind tests might be directed to particular analysts if the division of labor in the DNA unit of the laboratory permits it.

Logistically, the Random Reanalysis model is the least complicated of all blind testing models, when projected to a national scale. There is no manufacturing involved, and test performance "coordination" or oversight is much simpler. By our definition of external blind proficiency testing (§V.6 above), labs would have to outsource the entire audit and reanalysis. The costs associated with random reanalysis could be lower in terms of out-of-pocket expenses than with the other models. To the extent that laboratories cooperated with one another in providing random reanalysis testing and auditing for one another, the out of pocket expenses might be minimized (see §V.10 below).

3. Specification of Blind Test Modalities for a Large-Scale Program

If it is decided to make blind proficiency testing a requirement for the laboratories, the specific blind test models will also have to be considered. There are significant differences in logistics requirements, and therefore in the costs, of the different test models.

It will be important, therefore, to carefully consider what is wanted from a large-scale program before mandating it. As noted throughout the foregoing discussions in §V, somewhat different types of information can be obtained from different testing models. The following points should be considered in this context:

(i) Is it a program goal to generate a national picture of lab proficiency?
(ii) Is there to be national level review of the results?
(iii) Is "standardization" of proficiency tests, to make inter-laboratory results comparisons easier, a program goal?

The answers to these questions would focus the discussion, and probably determine which models of blind proficiency testing would be acceptable under the program.

J. Estimates of Costs of a Large-Scale Program

1. Fully Blind Test Models

The actual blind proficiency testing feasibility trials in this study were all external, done using what we have called the Blind / LE or Blind / CL models, and the test administration and manufacturing was centralized. Accordingly, we have cost data for this type of blind proficiency testing on a small scale, and can extrapolate those costs to a larger scale program based on certain assumptions. The costs for an external blind proficiency testing program following regional or local distribution / manufacturing models can only be estimated. Similarly, the costs of larger-scale programs based on Blind Analyst or Random Reanalysis models can only be estimated.

Based on our feasibility studies, it would require a minimum of two professionals to staff the test coordination / planning / administration office. For purposes of cost estimates, we will assume that one person is compensated at $50,000 per annum and the other at $35,000 for a total of $85,000 per annum (not including fringe benefits). We further estimate that the senior professional invests four person-days, and the junior professional invests three person-days for each test. Assuming 250 work days in a year, the personnel costs per test are thus: 4/250 x 50,000 + 3/250 x 35,000 = $1,220. Average travel to each test site is estimated to cost $1,200.

Costs of consumable supplies, communications, etc. are estimated at $200 per test. Manufacturing costs, based on our feasibility study experience, are $500 per test. Summing all the elements, the cost per test is $3,120.

Now assuming that 150 labs will be tested once annually, 150 tests have to be set up and run. With few or no economies of scale at the blind test set-up and implementation stage, the person-day requirements per test noted above could not be met by two people. 2.4 full time equivalent senior professional and 1.8 full-time equivalent junior professional people would be required. The costs of testing 150 labs annually under the assumption of $3,120 per test is $468,000. This estimate is unrealistically conservative. Allowing for a 20% fringe benefit rate, for example, the cost per test is $3,532. Rounding this number to $3,500, the cost of 150 tests is $525,000. Fringe benefits are included in all subsequent estimates, unless otherwise noted. In addition, costs associated with any national proficiency test oversight or review of results functions are not included. Including one meeting of a proficiency test oversight committee of 10 people at $1,000 per person adds $10,000 to the estimate and brings the total program costs to $535,000.

2. Economies of Scale

There are at least two ways to look at economies of scale in this context. First, it is likely that economies would be realized under a two test per lab per year requirement, i.e., it would probably not cost twice as much to do two tests per year as it did to do one per year. If we projected that the cost of doing two tests per year would be 150% of the one-test-per-year cost, the program would cost $794,700. Allowing for two meetings of 10 people at $1,000 each adds $20,000, and brings the total to $814,000. Second, we might project that costs would decrease over time if a large-scale program were implemented and sustained. This projection is based on a presumption that over time and with experience, it would become easier to set up tests with law enforcement agencies or conduit labs. Fewer person-days effort on the part of the test coordinators could be required, and travel would become less necessary. At the same time, it can be argued that these economies might not be realized because of the continuous changes in personnel assignments in law enforcement agencies. Thus, it could turn out to be just as complicated to set up tests in a third or fourth year as it was the first time, because the coordinators would be dealing with a new set of people. In addition, economies projected for future years would never be completely realized, because personnel and other costs would inevitably rise over time. For example, the one-test-per-year with one review committee meeting figure of $535,000 grows to over $650,000 after five years at an annualized growth rate of 5%.

3. Costs of a Program Under Blind Analyst Model

It is reasonable to assume that the costs of running a program under the Blind Analyst model would be close to the estimates for the "mature" fully blind program discussed in §V.J.2 above – about 57% of the estimated costs for the initial / startup phase of a Blind / LE, Blind / CL model program.

We assume here that, while laboratory administration is involved in planning the tests, there is still an external test coordinating entity tending to the details, manufacturing and transmittal to the labs.

4. Costs of a Program Under Random Reanalysis Model

The estimated costs of a program under the random reanalysis model assume that the entire reanalysis is conducted by an entity external to the laboratory (and external to the laboratory system, if applicable), to be consistent with our definition of an "external blind proficiency test" (§V.F above).

Here, the estimates are based on the assumption that the process would require 2 person days effort by an "auditor" (at $500 / day) and from 2 to 5 person-days effort by an "analyst" (at $350 / day). Included in the estimate are $1,200 travel costs and $200 consumables costs. The "auditor" has to visit the target lab to review candidate cases for reanalysis, choose one or more, then gather all the information, records, and evidence. The two to five person-day estimate for an analyst is based on the idea that it might take 2 person days effort to reanalyze a case involving multiplex-PCR and a series of STR locus typings, where it might take 5 person days effort to reanalyze a case involving a six-locus RFLP typing. These assumptions give a cost per test estimate in the range of $2,750 to $3,275. If the audit were conducted without reanalysis, the cost per test decreases to $2,200. For one audit per year (150), the cost is then $330,000 to $491,000. There would be almost no economies of scale in this model in projecting the costs from one per year to two per year, so the two-tests-per-year estimate would be about twice the one-per-year numbers.

These costs are roughly comparable to the estimates given for the Blind Analyst model, at least at the low end, where an audit does not include reanalysis of the biological evidence. At the high end, the costs are fairly comparable to those for the Blind / LE model. Here, although the costs are real, the actual out-of-pocket costs to laboratories might be lowered if laboratories could provide these services for one another. If some external organization or entity decided to offer these services to laboratories for a fee, however, the costs would then presumably be out-of-pocket to the target laboratories.

There is no requirement under this model for a national proficiency test-coordinating group, but there is likewise nothing that precludes having one.

5. Other Estimates of Costs

To try and provide a better picture of potential costs, we include estimates provided by both government agency and commercial provider representatives.

6. Cost Estimate Summary

A summary of the estimates presented in detail above, with explanatory footnotes as appropriate, is given in Table V-4 below. Costs have been rounded to the nearest thousand, except in the "cost / test" column.

Table V-4. Cost Estimate Summary

Blind Proficiency Test Program Model	Extrapolations from Present Project		
	Cost / Test	One Test Per Year Total	Two Tests Per Year Total
Blind / LE, Blind / CL	3,500	535,000 [a]	814,000 [b]
Blind Analyst	2,000	310,000 [a]	630,000 [c]
Random Reanalysis [d]	2,000 - 3,275	330,000 - 491,000	660,000 – 983,000
	Government Agency Provider Estimate		
Blind / LE, Blind / CL	10,000	1,510,000	3,020,000
	Commercial Proficiency Test Provider Estimate		
Blind / LE, Blind / CL	3,400	520,000	1,050,000
Blind Analyst	1,400	220,000	450,000

[a] Includes costs of one proficiency test review meeting

[b] 150% of one-test-per-year costs and includes two proficiency test review meetings

[c] Includes two proficiency test review meetings

[d] The low-end figure does not include reanalysis of the biological evidence

APPENDICES

A. TEXT OF THE DNA IDENTIFICATION ACT, AS ENACTED

Subtitle C-DNA Identification

SEC. 210301. SHORT TITLE.

This subtitle may be cited as the "DNA Identification Act of 1994".

SEC. 210302. FUNDING TO IMPROVE THE QUALITY AND AVAILABILITY OF DNA ANALYSES FOR LAW ENFORCEMENT IDENTIFICATION PURPOSES.

(a) Drug Control and System Improvement Grant Program. -Section 501(b) of title I of the Omnibus Crime Control and Safe Streets Act of 1968 (42 U.S.C. 3751(b)) as amended by section 150003, is amended-

(1) by striking "and" at the end of paragraph (23);

(2) by striking the period at the end of paragraph (24) and inserting "; and"; and

(3) by adding at the end the following new paragraph:

"(25) developing or improving in a forensic laboratory a capability to analyze deoxyribonucleic acid (hereinafter in this title referred to as 'DNA') for identification purposes."

(b) State Applications. -Section 503(a) of title I of the Omnibus Crime Control and Safe Streets Act of 1968 (42 U.S.C. 3753(a)) is amended by adding at the end the following new paragraph:

"(12) If any part of funds received from a grant made under this part is to be used to develop or improve a DNA analysis capability in a forensic laboratory, a certification that-

"(A) DNA analyses performed at such laboratory will satisfy or exceed then current standards for a quality assurance program for DNA analysis, issued by the Director of the Federal Bureau of Investigation under section 210303 of the DNA Identification Act of 1994;

"(B) DNA samples obtained by, and DNA analyses performed at, such laboratory will be accessible only-

"(I) to criminal justice agencies for law enforcement identification purposes;

"(ii) in judicial proceedings, if otherwise admissible pursuant to applicable statutes or rules;

"(iii) for criminal defense purposes, to a defendant, who shall have access to samples and analyses performed in connection with the case in which such defendant is charged; or

"(iv) if personally identifiable information is removed, for a population statistics database, for identification research and protocol development purposes, or for quality control purposes; and

"(C) such laboratory, and each analyst performing DNA analyses at such laboratory, will undergo, at regular intervals of not to exceed [*H8845] 180 days, external proficiency testing by a DNA proficiency testing program meeting the standards issued under section 210303 of the DNA Identification Act of 1994."

(c) DNA Identification Grants. -

(1) In general. -Title I of the Omnibus Crime Control and Safe Streets Act of 1968 (42 U.S.C. 3711 et seq.), as amended by section 210201(a), is amended-

(A) by redesignating part X as part Y;

(B) by redesignating section 2401 as section 2501; and

(C) by inserting after part W the following new part:

"PART X-DNA IDENTIFICATION GRANTS

"SEC. 2401. GRANT AUTHORIZATION.

"The Attorney General may make funds available under this part to States and units of local government, or combinations thereof, to carry out all or a substantial part of a program or project intended to develop or improve the capability to analyze deoxyribonucleic acid (referred to in this part as 'DNA') in a forensic laboratory.

"SEC. 2402. APPLICATIONS.

"To request a grant under this part, the chief executive officer of a State or unit of local government shall submit an application in such form as the Attorney General may require.

"SEC. 2403. APPLICATION REQUIREMENTS.

"No grant may be made under this part unless an application has been submitted to the Attorney General in which the applicant certifies that-

"(1) DNA analyses performed at the laboratory will satisfy or exceed then current standards for a quality assurance program for DNA analysis issued by the Director of the Federal Bureau of Investigation under section 210303 of the DNA Identification Act of 1994.

"(2) DNA samples obtained by and DNA analyses performed at the laboratory shall be made available only-

"(A) to criminal justice agencies for law enforcement identification purposes;

"(B) in judicial proceedings, if otherwise admissible pursuant to applicable statutes or rules;

"(C) for criminal defense purposes, to a defendant, who shall have access to samples and analyses performed in connection with the case in which the defendant is charged; or "(D) if personally identifiable information is removed, for a population statistics database, for identification research and protocol development purposes, or for quality control purposes; and

"(3) the laboratory and each analyst performing DNA analyses at the laboratory shall undergo, at regular intervals not exceeding 180 days, external proficiency testing by a DNA proficiency testing program that meets the standards issued under section 210303 of the DNA Identification Act of 1994.

"SEC. 2404. ADMINISTRATIVE PROVISIONS.

"(a) Regulation Authority. -The Attorney General may promulgate guidelines, regulations, and procedures, as necessary to carry out the purposes of this part, including limitations on the number of awards made during each fiscal year, the submission and review of applications, selection criteria, and the extension or continuation of awards.

"(b) Award Authority. -The Attorney General shall have final authority over all funds awarded under this part.

"(c) Technical Assistance. -To assist and measure the effectiveness and performance of programs and activities funded under this part, the Attorney General may provide technical assistance as required.

"SEC. 2405. RESTRICTIONS ON USE OF FUNDS.

"(a) Federal Share. -The Federal share of a grant, contract, or cooperative agreement made under this part may not exceed 75 percent of the total costs of the project described in the application submitted for the fiscal year for which the project receives assistance.

"(b) Administrative Costs. -A State or unit of local government may not use more than 10 percent of the funds it receives from this part for administrative expenses.

"SEC. 2406. REPORTS.

"(a) Reports to Attorney General. -Each State or unit of local government which receives a grant under this part shall submit to the Attorney General, for each year in which funds from a grant received under this part is expended, a report at such time and in such manner as the Attorney General may reasonably require which contains-

"(1) a summary of the activities carried out under the grant and an assessment of whether such activities are meeting the needs identified in the application submitted under section 2402; and
"(2) such other information as the Attorney General may require.

"(b) Reports to Congress. -Not later than 90 days after the end of each fiscal year for which grants are made under this part, the Attorney General shall submit to the Speaker of the House of Representatives and the President pro tempore of the Senate, a report that includes-

"(1) the aggregate amount of grants made under this part to each State or unit of local government for such fiscal year; and

"(2) a summary of the information provided in compliance with subsection (a)(1).

"SEC. 2407. EXPENDITURE RECORDS.

"(a) Records. -Each State or unit of local government which receives a grant under this part shall keep records as the Attorney General may require to facilitate an effective audit.

"(b) Access. -The Attorney General, the Comptroller General, or their designated agents shall have access, for the purpose of audit and examination, to any books, documents, and records of States and units of local government which receive grants made under this part if, in the opinion of the Attorney General, the Comptroller General, or their designated agents, such books, documents, and records are related to the receipt or use of any such grant."

(2) Table of contents. -The table of contents of title I of the Omnibus Crime Control and Safe Streets Act of 1968 (42 U.S.C. 3711 et seq.), as amended by section 210201(b), is amended by striking the matter relating to part X and inserting the following:

"Part X-DNA Identification Grants

"Sec. 2401. Grant authorization.

"Sec. 2402. Applications.

"Sec. 2403. Application requirements.

"Sec. 2404. Administrative provisions.

"Sec. 2405. Restrictions on use of funds.

"Sec. 2406. Reports.

"Sec. 2407. Expenditure records.

"Part Y-Transition-Effective Date-Repealer

"Sec. 2501. Continuation of rules, authorities, and proceedings."

(3) Authorization of appropriations. -Section 1001 of the Omnibus Crime Control and Safe Streets Act of 1968 (42 U.S.C. 3793), as amended by section 210201(c), is amended-

(A) in paragraph (3) by striking "and W" and inserting "W, and X"; and

(B) adding at the end the following new paragraph:

"(22) There are authorized to be appropriated to carry out part X-

"(1) $ 1,000,000 for fiscal year 1996;

"(2) $ 3,000,000 for fiscal year 1997;

"(3) $ 5,000,000 for fiscal year 1998;

"(4) $ 13,500,000 for fiscal year 1999; and "(5) $ 17,500,000 for fiscal year 2000."

(4) Effective date. -The amendments made by this section shall take effect on the date that is 60 days after the date of enactment of this Act.

SEC. 210303. QUALITY ASSURANCE AND PROFICIENCY TESTING STANDARDS.

(a) PUBLICATION OF QUALITY ASSURANCE AND PROFICIENCY TESTINGSTANDARDS. -(1)(A) Not later than 180 days after the date of enactment of this Act, the Director of the Federal Bureau of Investigation shall appoint an advisory board on DNA quality assurance methods from among nominations proposed by the head of the National Academy of Sciences and professional societies of crime laboratory officials.

(B) The advisory board shall include as members scientists from State, local, and private forensic laboratories, molecular geneticists and population geneticists not affiliated with a forensic laboratory, and a representative from the National Institute of Standards and Technology.

(C) The advisory board shall develop, and if appropriate, periodically revise, recommended standards for quality assurance, including standards for testing the proficiency of forensic laboratories, and forensic analysts, in conducting analyses of DNA.

(2) The Director of the Federal Bureau of Investigation, after taking into consideration such recommended standards, shall issue (and revise from time to time) standards for quality assurance, including standards for testing the proficiency of forensic laboratories, and forensic analysts, in conducting analyses of DNA.

(3) The standards described in paragraphs (1) and (2) shall specify criteria for quality assurance and proficiency tests to be applied to the various types of DNA analyses used by forensic laboratories. The standards shall also include a system for grading proficiency testing performance to determine whether a laboratory is performing acceptably.

(4) Until such time as the advisory board has made recommendations to the Director of the Federal Bureau of Investigation and the Director has acted upon those recommendations, the quality assurance guidelines adopted by the technical working group on DNA analysis methods shall be deemed the Director's standards for purposes of this section.

(b) ADMINISTRATION OF THE ADVISORY BOARD. - (1) For administrative purposes, the advisory board appointed under subsection (a) shall be considered an advisory board to the Director of the Federal Bureau of Investigation.

(2) Section 14 of the Federal Advisory Committee Act (5 U.S.C. App.) shall not apply with respect to the advisory board appointed under subsection (a).

(3) The DNA advisory board established under this section shall be separate and distinct from any other advisory board administered by the FBI, and is to be administered separately.

(4) The board shall cease to exist on the date 5 years after the initial appointments are made to the board, unless the existence of the board is extended by the Director of the Federal Bureau of Investigation.

(c) PROFICIENCY TESTING PROGRAM. -(1) Not later than 1 year after the effective date of this Act, the Director of the National Institute of Justice shall certify to the Committees on the Judiciary of the House and Senate that-

(A) the Institute has entered into a contract with, or made a grant to, an appropriate entity for establishing, or has taken other appropriate action to ensure that there is established, not later than 2 years after the date of enactment of this Act, a blind external proficiency testing program for DNA analyses, which shall be available to public and private laboratories performing forensic DNA analyses;

(B) a blind external proficiency testing program for DNA analyses is already readily available to public and private laboratories performing forensic DNA analyses; or

(C) it is not feasible to have blind external testing for DNA forensic analyses.

(2) As used in this subsection, the term "blind external proficiency test" means a test that is presented to a forensic laboratory through a second agency and appears to the analysts to involve routine evidence.

(3) Notwithstanding any other provision of law, the Attorney General shall make available to the Director of the National Institute of Justice during the first fiscal year in which funds are distributed under this subtitle up to $250,000 from the funds available under part X of Title I of the Omnibus Crime Control and Safe Streets Act of 1968 to carry out this subsection. [pH8846]

SEC. 210304. INDEX TO FACILITATE LAW ENFORCEMENT EXCHANGE OF DNA IDENTIFICATION INFORMATION.

(a) ESTABLISHMENT OF INDEX. -The Director of the Federal Bureau of Investigation may establish an index of-

(1) DNA identification records of persons convicted of crimes;

(2) analyses of DNA samples recovered from crime scenes; and

(3) analyses of DNA samples recovered from unidentified human remains.

(b) INFORMATION. -The index described in subsection (a) shall include only information on DNA identification records and DNA analyses that are-

(1) based on analyses performed by or on behalf of a criminal justice agency in accordance with publicly available standards that satisfy or exceed the guidelines for a quality assurance program for DNA analysis, issued by the Director of the Federal Bureau of Investigation under section 210303;

(2) prepared by laboratories, and DNA analysts, that undergo, at regular intervals of not to exceed 180 days, external proficiency testing by a DNA proficiency testing program meeting the standards issued under section 210303; and

(3) maintained by Federal, State, and local criminal justice agencies pursuant to rules that allow disclosure of stored DNA samples and DNA analyses only-

(A) to criminal justice agencies for law enforcement identification purposes;

(B) in judicial proceedings, if otherwise admissible pursuant to applicable statutes or rules;

(C) for criminal defense purposes, to a defendant, who shall have access to samples and analyses performed in connection with the case in which such defendant is charged; or

(D) if personally identifiable information is removed, for a population statistics database, for identification research and protocol development purposes, or for quality control purposes.

(c) FAILURE TO COMPLY. -Access to the index established by this section is subject to cancellation if the quality control and privacy requirements described in subsection (b) are not met.

SEC. 210305. FEDERAL BUREAU OF INVESTIGATION.

(a) PROFICIENCY TESTING REQUIREMENTS. -

(1) GENERALLY. -(A) Personnel at the Federal Bureau of Investigation who perform DNA analyses shall undergo, at regular intervals of not to exceed 180 days, external proficiency testing by a DNA proficiency testing program meeting the standards issued under section 210303

(B) Within 1 year after the date of enactment of this Act, the Director of the Federal Bureau of Investigation shall arrange for periodic blind external tests to determine the proficiency of DNA analysis performed at the Federal Bureau of Investigation laboratory.

(C) In this paragraph, "blind external test" means a test that is presented to the laboratory through a second agency and appears to the analysts to involve routine evidence

(2) REPORT. -For 5 years after the date of enactment of this Act, the Director of the Federal Bureau of Investigation shall submit to the Committees on the Judiciary of the House and Senate an annual report on the results of each of the tests described in paragraph (1).

(b) PRIVACY PROTECTION STANDARDS. -

(1) GENERALLY. -Except as provided in paragraph (2), the results of DNA tests performed for a Federal law enforcement agency for law enforcement purposes may be disclosed only-

(A) to criminal justice agencies for law enforcement identification purposes;

(B) in judicial proceedings, if otherwise admissible pursuant to applicable statutes or rules; and

(C) for criminal defense purposes, to a defendant, who shall have access to samples and analyses performed in connection with the case in which such defendant is charged.

(2) EXCEPTION. -If personally identifiable information is removed, test results may be disclosed for a population statistics database, for identification research and protocol development purposes, or for quality control purposes.

(c) CRIMINAL PENALTY. -(1) A person who-

(A) by virtue of employment or official position, has possession of, or access to, individually identifiable DNA information indexed in a database created or maintained by any Federal law enforcement agency; and

(B) knowingly discloses such information in any manner to any person or agency not authorized to receive it, shall be fined not more than $100,000.

(2) A person who, without authorization, knowingly obtains DNA samples or individually identifiable DNA information indexed in a database created or maintained by any Federal law enforcement agency shall be fined not more than $100,000.

SEC. 210306. AUTHORIZATION OF APPROPRIATIONS.

There are authorized to be appropriated to the Federal Bureau of Investigation to carry out sections 210303, 210304, and 210305-

(1) $5,500,000 for fiscal year 1996;

(2) $8,000,000 for fiscal year 1997;

(3) $8,000,000 for fiscal year 1998;

(4) $2,500,000 for fiscal year 1999; and

(5) $1,000,000 for fiscal year 2000.

B. NATIONAL FORENSIC DNA REVIEW PANEL

Chairman:
Joshua Lederberg, Ph.D.
Professor
The Rockefeller University
New York, NY

Shirley S. Abrahamson, J.D., S.J.D.
Chief Justice
Wisconsin Supreme Court
Madison, WI

Jack Ballantyne
Associate Professor
University of Central Florida
Department of Chemistry
Orlando, FL

Bruce Budowle, Ph.D.
DNA Laboratory Manager
Forensic Science Research
and Training Center
FBI Academy
Quantico, VA

Ranajit Chakraborty, Ph.D.
Professor
University of Texas at Houston
Department of Human Genetics
Houston, TX

Benard Devlin, Ph.D.
Professor
Carnegie Mellon University
Department of Statistics
Pittsburgh, PA

Arthur Eisenberg, Ph.D.
Director of DNA Laboratory
University of North Texas
Health Science Center
Ft. Worth, TX

Project Monitor:
Richard Rau, Ph.D.
Forensic Science Program Manager
National Institute of Justice
Washington, DC

Marcia Eisenberg, Ph.D.
Director of Forensic Identity Laboratory
and Research & Development
Laboratory Corporation of America
Research Triangle Park, NC

Paul B. Ferrara, Ph.D.
Director
Virginia Division of Forensic Sciences
Richmond, VA

Honorable Susan Gaertner
Ramsey County Attorney
St. Paul, MN

John W. Hicks
Assistant Director
Alabama Department of Forensic Sciences
Birmingham, AL

Margaret Kuo
Chief Criminalist
Orange County Crime Laboratory
Santa Ana, CA

Terry L. Laber
Forensic Sciences Supervisor
Minnesota Department of Public Safety
St. Paul, MN

Randall S. Murch, Ph.D.
Deputy Assistant Director
Federal Bureau of Investigation
Laboratory Division
Washington, DC

Dennis Reeder, Ph.D.
Group Leader
National Institute of Standards
and Technology
DNA Technology Group /
Biotechnology Division
Gaithersburg, MD

Moses Schanfield, Ph.D.
Director
Analytical Genetic Testing Center
Denver, CO

Barry C. Scheck, Esq.
Professor
Benjamin N. Cardozo School of Law
New York, NY

William C. Thompson, Ph.D.
Professor
University of California - Irvine
School of Social Ecology
Irvine, CA

Victor W. Weedn, M.D., J.D.
Laboratory Director
Alabama Department of Forensic Sciences
Birmingham Regional Laboratory
Birmingham, AL

David J. Werrett, Ph.D.
Director of Research and DNA Services
The Forensic Science Service
Birmingham, United Kingdom

C-1. SUMMARY OF THE FIRST NATIONAL FORENSIC DNA REVIEW PANEL (NFDRP) MEETING

February 20, 1997
9:00 am - 12:30 pm
New York, NY

NFDRP Members Present: Jack Ballantyne, Bernard Devlin, Dennis Reeder, Margaret Kuo, Arthur Eisenberg, Marcia Eisenberg, Bruce Budowle, Joshua Lederberg (Chair), John Hicks, Shirley Abrahamson, Paul Ferrara, Moses Schanfield, William Thompson, David Werrett, Randall Murch, Terry Laber, Barry Scheck, Susan Gaertner

NFDRP Members Absent: Victor Weedn, Ranajit Chakraborty

NIJ Grant Monitor: Richard Rau

Project Staff: Joseph Peterson, R.E. Gaensslen, and George Lin

Prior to the beginning of the committee meeting, project staff distributed three-ring binders to members containing the meeting agenda, NFDRP membership list, project summary, a draft proficiency testing literature review, the laboratory questionnaire, agency participation agreement, donor consent form, and travel reimbursement form.

Dr. Lederberg called the meeting to order. He asked that all members introduce themselves and to indicate if they had any conflicts of interest they wished to declare to other committee members.

Next, the NIJ grant monitor, Dr. Richard Rau, charged the panel:

1. Determine whether a double blind external DNA proficiency testing program is feasible for public and private laboratories performing DNA analysis.

2. Recommend a National Model Forensic DNA External Proficiency Testing Program for public and private laboratories performing DNA analysis.

3. Deliberate and provide direction to the project staff in order to ensure that all issues and avenues of approach are adequately investigated and evaluated.

Following Dr. Rau's charge to the panel, Drs. Peterson and Gaensslen made a slide presentation describing the goals of the project and progress of the study to date. The two-year project grant from NIJ was signed on May 31, 1996 and project activities began in June. Project objectives include: establishment of the NFDRP; review of proficiency testing literature; survey of forensic DNA testing laboratories/relevant law enforcement agencies; design of alternative blind proficiency testing approaches; limited feasibility study of blind testing; and report/recommendations to the NFDRP.

Review of proficiency testing (PT) literature primarily focused on proficiency testing in the clinical laboratory setting. The findings and current practices in the clinical laboratory were presented, including an overview of clinical laboratory regulation. In addition, the few studies that examined blind (undeclared) PT vs. open (declared) PT results were highlighted. The more limited forensic DNA proficiency testing literature was reviewed as well. Committee members were encouraged to review the draft and to respond with corrections and suggestions.

The preliminary data from the survey of crime laboratories was presented via the slides. As of January 1997, approximately 25%(44) of the 150 surveys mailed to laboratories had been returned. Consequently, the data discussed at the meeting were very preliminary in nature. Data presented included: laboratory affiliation, size of laboratory in terms of analysts, status of laboratory certification, types of DNA analyses being performed, quality assurance/quality control procedures being followed, and number of proficiency tests taken per year. Also discussed were replies indicating alternatives to blind proficiency testing, including random re-analysis, being practiced in laboratories.

Discussion

The discussion began with clarification of the meaning of "blind external proficiency testing." The term "double blind" is used in studies evaluating the effects of medication where neither the physician nor the patient is aware if they are administering/receiving the experimental or control drug. Such a model does not apply to proficiency testing in a forensic laboratory, and consequently the term "double blind" is not used in the DNA Identification Act itself. The DNA Act defines, "'blind external proficiency test' as a test that is presented to a forensic laboratory through a second agency and appears to the analyst to involve routine evidence." The terms "external" and "second agency" indicate the test would be originating from an organization outside the laboratory. It was noted that ASCLD-LAB guidelines define "external" as outside the agency. It is not clear as to whether samples originating from other laboratories in a statewide system might satisfy the external, second agency requirement. Nor is the DNA Act clear if "blind" refers only to the specific examiner or that the entire laboratory hierarchy is blinded to the fact that the sample is a test. Mr. Hicks and others indicated project staff might be able to clarify the meaning of this language by researching the legislative history of the Act.

Project staff indicated that they will review various "levels of blindness" – that is, where only the examiner is blinded to the fact the evidence is a test, to where all examiners and perhaps the supervisor in the DNA testing section are blinded, to a situation where everyone associated with the laboratory is blinded, including the laboratory director. The initial feasibility test will attempt to keep everyone associated with the laboratory blinded; this may be adjusted, however, in subsequent examinations. Obviously, the more persons to be blinded, the more intricate (and costly) the test administration becomes. The point was also raised as to how one can be certain that the examiner truly did not detect the blind sample as a test. No solution to this question was given.

Internal Proficiency Testing - During the slide presentation, a question was asked regarding one of the questions on the survey concerning internal proficiency tests. Some laboratories think that running internal lane standards and random samples in gels, where the correct results are not known to the examiner in advance, as an internal PT program. Some members agreed that this was a form of internal PT. Furthermore, the question was asked, "Can this activity be considered a viable proficiency testing paradigm for forensic laboratories?" Another committee person made the point that these internal lane standards/random samples are not proficiency tests, rather they are more appropriately described as QC

checks. An additional question is if laboratories were to maintain records of the success of examiners in running these internal controls, are these data meaningful measures of laboratory proficiency?

Commentary on Blind Proficiency Testing - The goals and objectives of a blind proficiency testing program were discussed. These include: (1) to determine which, if any, analysts are doing substandard work; (2) to determine biases, if any, in the laboratory; and (3) to clarify potential problems.

The following issues and problems associated with blind proficiency testing were discussed:

- large dedicated staff would be required to operate a national program
- because blind tests would be treated as actual evidence, it is not possible to designate which scientist receives the test, and when the results will be returned
- such a program would be expensive
- possibility of experienced analysts recognizing the samples as part of a proficiency test
- the need for tests to be suitably complex and realistic (if not, then what's the purpose?)
- the more complex the manufacturer attempts to make a test the more opportunities for a slip up and recognition by the analyst the case is really a blind PT. Plus, there are more chances for problems to arise among multiple labs, i.e., more difficulties in getting a consensus value
- the need for uniformity of proficiency tests on one hand, but on the other the tests must be ultimately tailored to the many different criteria employed by laboratories in processing and examining specimens
- will only analytical results be reviewed, or will the entire evidence testing process from receipt of the evidence to the final report - be examined

Because one of the goals of this research is to conduct a limited feasibility study of blind proficiency testing, there will be the need to approximate the costs of such testing. It was mentioned that realistically the study cannot be that quantitative, but nonetheless should examine costs, logistics, and overall benefits. The committee recommended that a model should be developed which every jurisdiction will be able to utilize.

Donor Consent - The issue of donor consent was discussed. The objective of the limited feasibility study will be to prepare realistic biological evidence items, such as would be encountered in routine crime laboratory casework. As such, biological specimens are required from donors. The primary concern is that the donor's DNA profile will be entered into a database for period of time, from a few days up to several months. If a match should occur with an actual case, the police might want to seek the real identity of the donor. Although a condition for participation is that the laboratories must agree to remove databank records of PT samples, the researchers cannot guarantee to the donor this will occur (see draft donor consent form in binder).

The committee seemed comfortable with procedures to protect the interests/identity of the donor, particularly since a fictitious name will be used. One highly unlikely, but nevertheless possible, eventuality is for the donor sample to be matched with another file in the data bank. The project staff mentioned that if there was a match, law enforcement officials might attempt to determine the donor's true identity by questioning the person who collected the samples and investigate their records. It was suggested that the project staff investigate alternative coding systems that may be available.

The committee suggested the researchers investigate a "waiver of indemnification" that could be added to the donor consent form, in which the donor would hold the University of Illinois at Chicago harmless.

Matches in DNA Indexes -The next topic of discussion concerned problems that could arise if the blind test leads to databank or CODIS matches. If several blind proficiency tests utilizing the same donor specimens are distributed to different laboratories in different states, one or more "hits" may result. This would potentially alert the laboratory and others who have tested the samples that the specimens were indeed proficiency test specimens. One member noted that this is not necessarily a bad thing because such a "hit" might actually serve as a good check to determine if the data banking system is working properly. (It is was noted that the FBI is currently working on the guidelines and procedures to be followed by national CODIS and the involved participant laboratories specifically in the event a match occurs between multiple states.)

Purging of Documents and Records of a Proficiency Test - A proficiency test that has been worked through a laboratory will have associated with it a series of documents, simulated evidence, worksheets, and analytical reports in addition to the profiles entered into a DNA databank. Originally, the project staff believed laboratories should make every effort to purge all records and data from the proficiency test. The committee members, in general, felt this was neither possible nor necessarily desirable. The chairman expressed concern that purging such records, per se, might be considered "tampering with a good system." Maintaining the integrity of a crime laboratory's chain of custody/record keeping system is extremely important to the quality of work, therefore removing records from such a system could actually create new additional problems. The committee members agreed that at the very least, the DNA type entered into the computer database should be removed. All other documents, records, and case notes should probably be retained and identified as part of a proficiency test. These records would in fact document the laboratory's participation in proficiency testing as well as its performance.

Deception - In order for an external blind proficiency test to be successful, a fictitious case must be fabricated along with all the accompanying documents and records. Because the submitting agency (e.g., prosecutor or law enforcement) must submit the "case" to the laboratory, it plays a role in deceiving the laboratory. At least one committee member thought this issue of deception repugnant. Some laboratories in their response to the survey feared the damage such a practice might have on the trusting relationship between laboratory and law enforcement agency. While some committee members shared this concern and saw the investigating officer and laboratory analyst as members of a team, others believed that laboratories and user agencies needed to be more independent. With such independence, the use of blind proficiency testing and deception should be less of a problem. The point was made that scientists and police officers needed to be educated on the benefits of a blind proficiency program. The fact that deception occurs is incidental, because such a program would be for the greater good of the system.

It was also mentioned that there was considerable precedent for employing deception in examining the operations of criminal justice agencies, as where integrity officers within police agencies (such as New

York City) have the duty to test the integrity of law enforcement personnel as well as expose corruption, from the patrol officer up to the level of the judiciary. It may ultimately come down to a balancing of the net benefits to be derived from a program like blind testing that would involve some deception, versus the costs (monetary and other) to the laboratories and the criminal justice system

It is worth noting that some sections of a forensic laboratory may be more independent than others. For example, some laboratories utilize a criminalist who examines evidence in the context of the complete case. Specimens to be DNA tested are selected by the criminalist and sent to the DNA section of the laboratory. The analyst may know nothing about the case itself, but tests all specimens received from the criminalist. It was mentioned that blind testing of the analyst under this model would be straight forward, while testing of the criminalist more of a challenge.

Alternatives to Blind Proficiency Testing - Alternatives to blind proficiency testing were discussed. The committee chairman suggested that alternative approaches should be weighed. In particular, random re-analysis was discussed. He framed the question "What problem does proficiency testing solve that cannot be addressed in any other way?" He stated he had heard few arguments that convinced him blind proficiency testing was better than random re-analysis. Random re-analysis is a procedure in which specimens from a randomly selected case, previously examined, are re-analyzed. Because laboratory analysts do not know which of their cases may subsequently be chosen for re-analysis, they were technically blinded to it being a test. Some committee members were enthusiastic about this method because it permits a complete examination of the receipt, testing and reporting of results, and it is much less expensive to administer. Depending upon how thorough this process is, random reanalysis can check the analytical results as well as the evidence handling decisions. If, however, the re-analysis only re-tests what the initial examiner tested then it is conceivable the re-analysis would not detect original errors where evidence stains, for example, were missed/omitted.

Final Comments - The committee chairman made the point that PT should not only look at technical aspects of DNA testing, but also look at all the criteria for a case to get worked. Furthermore, he reiterated that a future study might be necessary to examine the whole process by which evidence is collected and transferred to the forensic laboratory for DNA testing.

Dr. Werrett reported on the results of the blind DNA proficiency testing program in the United Kingdom. In this program, not only is the DNA section tested, but all other sections of the laboratory (e.g., toxicology, trace, etc.) as well. Annually, each section of the laboratory meets with the Quality Assurance Unit (which sets up the blind proficiency test) to determine and agree upon what types of materials are to be included in blind proficiency tests. These proficiency tests are delivered by police officers as real cases sometime during the year. This blind proficiency testing program has been conducted since 1977, with blind DNA testing beginning in 1987. Dr. Werrett also noted that if you wish to look at cases in their entirety, then blind testing is the preferred method of approach; if one only wants to look at the performance of specific DNA tests, then declared proficiency testing may be just as good.

The project staff will be in contact with committee members in setting the date for the next advisory meeting. It was felt that the results of the initial feasibility tests need to be available before a next meeting is scheduled.

The meeting adjourned at approximately 12:30 pm.

C-2. SUMMARY OF THE SECOND NATIONAL FORENSIC DNA REVIEW PANEL (NFDRP) MEETING

Sunday December 7, 1997
9:00 am - 1:00 pm
Chicago, IL

NFDRP Members Present: Joshua Lederberg (Chair), Jack Ballantyne, Dennis Reeder, Margaret Kuo, Arthur Eisenberg, Moses Schanfield, William Thompson, David Werrett, Terry Laber, Barry Scheck, Susan Gaertner, Victor Weedn, Ranajit Chakraborty

NFDRP Members Absent: Bruce Budowle, John Hicks, Randall Murch, Shirley Abrahamson, Paul Ferrara, Bernard Devlin, Marcia Eisenberg

NIJ Grant Monitor: Richard Rau

Project Staff: Joseph Peterson, R.E. Gaensslen, and George Lin

Prior to the meeting, project staff distributed a draft summary of the survey data, and meeting agenda to members.

Dr. Lederberg called the meeting to order. Dr. Rau introduced a guest, Mr. Christopher Asplen, Assistant U.S. Attorney who described the creation of the National Commission on the Future of Forensic DNA Evidence. This is a two-year commission with the group's objective to determine policies that will maximize the value of DNA in the criminal justice system. Among the issues to be addressed by this commission will be post conviction matters, and other legal issues, laboratory funding, crime scene/evidence gathering procedures, and research/technology alternatives. Mr. Jeremy Travis, Director of the National Institute of Justice, will be appointing members of the commission to include laboratory scientists, lawyers, law enforcement personnel, medical examiners, and ethicists. It was stressed that the group will not duplicate the work of the DAB and the NFDRP.

Following this announcement, Drs. Peterson and Gaensslen made a slide presentation reviewing project goals, research plan and progress of the study to date. The emphasis of the presentation was (1) survey results of DNA laboratories and law enforcement agencies and (2) feasibility testing of blind proficiency tests. Dr. Peterson highlighted the survey data described beforehand in the draft summary. Dr. Chakraborty suggested that we further examine the types of laboratory respondents that answered various questions. (That is, were some types of laboratories over or under represented in responses to particular questions?) Dr. Gaensslen's presentation covered various feasibility testing topics, including: (a) agreements with testing laboratories, law enforcement agencies, and conduit laboratories (b) arrangements with the manufacturer of proficiency tests; (c) donor recruitment and informed consent provisions (d) manufacture of blind cases and evidentiary items; (e) demographics and description of target laboratories; (f) results of tests returned and cases cleared; (g) NY State blind testing program results; (h) comparison of results of blind tests with reference lab values; (I) estimated costs of blind tests; and (j) possible models for a large-scale program and estimated costs.

Dr. Gaensslen described the types of laboratories and cases submitted in the field-testing. Whereas the project is consider several alternative blind PT approaches, the project staff decided to field test 1) completely blind and 2) blind/conduit laboratory approaches. To facilitate cooperation, the project staff made visits to four locales/police agencies that agreed to serve as agencies submitting blind PTs to various laboratories. Eight proficiency tests were manufactured and issued to 2 private, 5 state, and 1 local laboratory. Seven rape and one blood transfer scenario were employed in the field test. As of the December NFDRP meeting, the project staff had received results from two laboratories.

Discussion

Internal Proficiency Testing - While reviewing the survey data, there was discussion as to what is considered an internal PT. For example, would QC results on built-in internal lane controls be considered an internal PT? These controls are typically run in every gel and the results unknown to the analyst. Although PT is considered a QA procedure, the distinction was made that internal lane controls are used for continuous quality monitoring while proficiency tests are for competency assessment. Proficiency testing may also be defined broadly to encompass the complete case including handling of evidence, isolation of stains, extraction of DNA, interpretation of results, report writing, clerical, and chain of custody. There was general agreement among panel members that internal lane controls would not be considered internal PTs.

The point was brought up that ASCLD/LAB requires two DNA proficiency tests for each examiner per year, one of which must be external. Most laboratories are opting to take two external proficiency tests because they are much easier to administer than an internal PT. Smaller laboratories, in particular, have fewer personnel and resources to set up internal PTs.

New York State Experience in Blind PT - Following Dr. Gaensslen's slide presentation, Mr. Scheck related the experiences of a blind PT feasibility study being performed in New York state. He mentioned that New York passed legislation to form the Forensic Science Review Commission, out of which a DNA Advisory Board was created. The Forensic Science Review Commission sets standards, regulates forensic laboratories in the state, and has statutory authority over them. A subcommittee of the DNA Advisory Board commissioned a feasibility study of blind PT which was designed by Carl Selavka and Paul Scheckman. There are five laboratories performing forensic DNA testing in the state and they have now completed testing of one laboratory. Because the Board holds regulatory authority over all crime laboratories in the State, they do not need the laboratories' consent to engage in this type of testing.

Mr. Scheck noted key points in the design are: (1) limiting the number of persons involved; (2) securing the cooperation of the principal law enforcement authority in the pertinent jurisdiction; usually this is district attorney or detective supervisor who initiates the case; (3) scientists and reference laboratory design the test; (4) turnaround time for the first test was 45 days. The costs have not yet been determined and laboratories have not been notified that they were tested. Mr. Scheck mentioned that the design of the New York study is similar to UIC's in that no one in the target laboratory was aware of the particular test and that cooperation of external (law enforcement) agency was secured.

Purpose of Proficiency Testing - On the purpose of blind proficiency testing, the panel agreed that blind PTs "make laboratories stay on their toes." An opinion given was that blind PTs essentially test for analyst "trustworthiness" (e.g., verifying are analysts are following laboratory guidelines?), should test the total system, and should inspire public confidence in the laboratory. Some members felt that blind PT

is another QA practice that adds value and credibility to the testimony of the analyst or laboratory in the courtroom. One panelist noted that blind PT may be beneficial, but not really necessary, because of all the other QA/QC procedures in place in the laboratory (e.g., training, certification of analysts, etc.), which serves a similar purpose. It was made clear that blind PT is not a "magic bullet"--that it is only a snapshot of the laboratory's work at a particular time. Blind PTs may not find all errors in the laboratory, but for that matter, there are no QA/QC measures that will find every error. It is also worth mentioning blind PTs are better able to examine the non-analytical phase of testing while open PTs focus more on the analytical phase of the testing process.

One important area that the panel discussed was whether blind proficiency tests are designed to test the laboratory (the total system) or the individual analyst. Generally, the structure of blind tests makes it very difficult to target specific analysts. If the goal was to blind test an analyst a certain number of times per year, it would be difficult, particularly in larger laboratories where there are many analysts. If incoming cases are assigned randomly, many PTs might have to be introduced before all analysts were tested. Problems also arise when analysts specialize. (i.e. in PCR, RFLP, and STRs). A representative from a private laboratory noted that it was sometimes possible to channel a PT to a specific analyst as where the submitter has experience with a particular examiner and requests that he/she handles the case. The seriousness of the case also may determine who works a case. If the case is more serious and/or difficult, the laboratory may assign an analyst with more experience to analyze the case. The consensus of the panel is that, essentially, blind tests tend to test the entire system rather than individual analysts.

Dr. Weedn made three points regarding PTs: (1) errors found by PTs are typically non-recurring errors--that is, after an error is detected, corrective actions are usually taken, thus ensuring the error does not recur; (2) PTs have no built-in redundancies like casework (for example, test results of multiple specimens in casework allows the analyst to verify his work); and (3) PTs generally take on two forms--one form is that they are designed to regulate laboratories, and the other is for educational purposes. The primary distinction between these PTs is that, generally, educational PTs do not penalize the analyst or laboratory, if an error is made. In addition, PTs designed to regulate are not ordinarily challenging and are designed to confirm that laboratories meet minimum regulatory standards. Finally, Dr. Weedn mentioned that the College of American Pathologists believes there is greater utility in educational PTs than regulating PTs in terms of laboratory improvement.

Expected Performance of Laboratories - The panel discussed results of proficiency testing and what is considered to be correct and incorrect results. There are discretely right and wrong results, such as DNA types, and there are also results that are more subjective such as an accepted DNA extraction efficiency (x%), typing of mixed stains, or typing of marginal samples. It was stated by Dr. Werrett that the expected performance of a laboratory on a test must be defined prior to blind testing. Dr. Werrett asked if we have prepared a written statement describing expected laboratory performance for the tests we have issued. Because this project is a feasibility study, expected performance was not specified. Dr. Eisenberg also asked about the language of the DNA Act and if it included the grading of PT. The DNA Act does not describe grading of blind PTs; however, Dr. Rau suggested that the project staff consider evaluating such a procedure and possibly analyze the test results of this study.

The panel discussed circumstances in which errors occur in the laboratory. It was mentioned that more errors occur when samples are marginal. A NIST study conducted by Dr. Reeder addressed this issue by using 30 ng mixed samples that were tested by various laboratories. NIST determined that there was a 75% success rate with 25% getting partial to no results. These errors were most often caused by

transcription mistakes and DNA extraction inefficiencies. It was observed that more experienced laboratories tended to do better.

One panelist noted that one of the major problems that would be addressed by blind PTs is errors of omission, not commission. The example discussed was the selection of stains from garments or other evidentiary items for DNA typing. Dr. Gaensslen described a scenario involving six stains, five from the same donor, and the sixth from someone else. In a case involving a multiple blood stained sock, what would be the "correct" number of stains to sample and which should be typed? Presumably, the laboratory has a protocol that analysts should follow in a more "complex" case. Dr. Werrett pointed out that the subjectivity of these errors are really differences in the interpretation of laboratory evidence handling policies rather than analytical procedures.

During this discussion, Dr. Lederberg also suggested that these issues (e.g., DNA extraction efficiency, mixed stain analysis, sampling, etc.) should be studied to gain a better understanding of the frequency and reasons these errors occur. One member was concerned about analysts missing small stains and argued that blind PT would address this situation. That is, if blind PTs show examiners miss the one stain that excludes the suspect, then this may persuade labs and examiners to review their evidence more carefully and test more exhaustively. Some members agreed that so-called educational PTs would be a good method to study this question. Others believed that using PT in such a fashion is not really an efficient way to study this issue ("...it is cumbersome, costly, and sporadic"). Others suggested that these problem areas could be elucidated by using other types of measures such as training, reviews, QC samples, and "challenge" testing. Challenge testing involves repeated testing focusing on a specific problem area such as extraction efficiency. It was stressed, once again, that challenge testing should be of reasonable difficulty.

One panel member's opinion is that the major problem in laboratories is not the way analysts run the gels, but how the evidence is selected and handled by police and laboratory scientists. There is a need to check the forensic training of analysts, not just their scientific skills. Many concur it is more basic criminalistics skills that are deficient among some DNA analysts. The question, therefore, is if blind PT is the best way to address this problem. Furthermore, if the problem is at the crime scene and procedures followed to identify and recover evidence rather than in the laboratory itself, why invest more funding in the laboratory area?

Design and Complexity of PT - The matter of how challenging PTs should be was discussed. Difficult PTs will help to identify the problem areas in a laboratory. Ms. Gaertner noted that if blind PT is not getting at the truly marginal/difficult cases, and then is it really worth the effort? A blind PT could be designed to be more challenging by employing marginal samples, complex blood splatter, and case scenarios. The introduction of samples of smaller size, including blanks, where contamination conceivably could be introduced by the laboratory, would better simulate actual casework. Another challenging model would be to set up a test where PCR testing only would include a suspect, but RFLP would exclude the suspect. Finally, if the analytical phase is far less of a problem than the evidence handling phases of a case, then PT should incorporate these stages into the process--otherwise PT would be avoiding the real problems. Mr. Scheck believes these types of cases would be ideal blind PTs.

One caveat in designing a PT is the "reasonableness" of the test. Depending on the number of stains and complexity of the case, it is a possibility that an analyst could miss a small stain and "fail" the PT. The panel agreed that attention must be paid to these criteria before the development and distribution of the blind PT.

Costs of a Blind PT - The project staff's initial estimates from the current study are that the cost of a blind proficiency test may be in excess of $2000/case. A question that must be addressed is whether the costs of a blind proficiency test are justified by the benefits. Dr. Lederberg noted that we may have already shown that blind PT is feasible, since at least one laboratory performed the test without detection. But we must also consider the costs of blind PT, both in terms of manufacturing and issuing the samples, but also in terms of the costs to the participating laboratories. We also have to consider what is accomplished with blind PT and the relative merits of blind PT and other alternatives. (That is, one of the proposed virtues of blind PT is that it keeps the examiner on his/her toes and they never know which case might be a test). The questions are:

1) What are the objectives of blind PT?
2) Are there any other QA/QC measures that achieve the same objectives as blind PT?
3) If there are other alternatives, are they more cost effective than blind PT?

In sum, Dr. Lederberg expressed his belief that the panel needed to determine: (1) the fiscal costs of a blind PT; (2) what is accomplished from blind PT; and (3) the cost-effectiveness of alternative modalities.

Ms. Gaertner commented on the fact that our cost estimates thus far do not include the costs of law enforcement agency personnel and prosecutor involvement. Ms. Gaertner stated that this expense may loom as a very important cost consideration. Dr. Gaensslen noted that he thought our law enforcement contacts spent about ½ to 3/4 of a day on the case. In addition, Dr. Gaensslen noted that not only do the primary law enforcement case principals need to be briefed about the PT case, but also a back-up person in case the primary individual is absent when the laboratory makes an inquiry. (It is noted here that Dr. Rau believed that the law enforcement position needs to be represented on this panel's deliberations.) On the matter of costs, Mr. Ballantyne noted an important consideration, that being the peer review of blind PT laboratory reports after the testing is completed. Another cost issue will be that of reference laboratories, especially if PT tests contain specimens from more than a single donor. Mr. Scheck noted that the increasing use of STRs should lower costs (presumably, because they are automated). Dr. Weedn noted another important cost issue which is the organization of the blind PT distribution network. In a centralized model where there is only one source of PTs, presumably, costs would decrease due to greater economies of scale. Whereas, in a regionalized model where there are multiple sources, the cost of a PT would increase. Comment was made that the review process might even be more costly and lengthy than the actual manufacture and distribution of the test. (It must also be decided who is qualified to perform the review.) Another area that impacts cost is related to the complexity of the case--presumably the more complicated/challenging the PT, the more costly it will be to implement and evaluate.

Given the fact that blind PT may be more costly than alternative procedures, the benefits of blind PT must be carefully determined. Mr. Scheck believes there is great symbolic value for the prosecution to be able to report their laboratory passed a blind test. Still, others believe that blind PT, regardless of cost, is absolutely necessary and that all possible forms of QA/QC must be employed to minimize errors. Mr. Sheck notes that 85% to 90% of cases are handled by public defenders and they receive a pittance to finance scientific reviews and/or re-analysis of prosecution evidence. He notes this funding issue is an intractable legal problem. Therefore, maximum effort must be made to strengthen the professionalism of laboratories and their QA/QC work.

Dr. Werrett reported that the English Forensic Science Service performs blind PTs and recognizes the high costs of setting up cases. The Forensic Science Service spends 20-30% of its time checking the reports and data of blind PTs. Although costly, the English believe it is worth the expense. He also mentioned that the principal value of blind PT is not the results per se but efforts to try to understand the reasons for errors and the institution of remedial programs.

Alternative QA/QC Methods - There are many different ways to assess performance, but the question remains, "What is the most cost effective way to achieve this?" Audits, random reanalysis and/or other modalities might achieve the same ends as blind PT and at a lesser cost. Auditing and re-testing were discussed as alternative methods to blind PT--either used in conjunction with PT or as substitutes. An audit is considered to be an external review of a case. The entire case is comprehensively reviewed including chain of custody, analytical, and non-analytical phases of the case. If any problems are discovered in the audit, the case may be re-worked. Some members feel that audits are the best mechanism for checking chain of custody because all results and documentation are checked, which gives a more complete picture. Mr. Scheck noted that one problem of the audit is that it can only review closed cases, and with the prospect of appeals, may delay for years the possibility a given case will be eligible for review. One member also noted that one may anticipate pressures on auditors if a particularly "big" case is selected for review. In addition, to retest evidence in an actual criminal case, one may also need the consent of various parties. Dr. Werrett notes that cultures form in laboratories for doing things in certain ways and that it is absolutely essential that audits must be performed by outsiders. Dr. Werrett noted another problem with auditing which is the fact that there is not an agreed upon set of answers (outcome) for the analysis before the tests are performed. In contrast, with blind PT, one decides beforehand what the outcome should be.

Re-testing/re-analysis was another method discussed. While this is another type of competency assessment, this method requires that there is enough sample left over to be tested. One of the major problems is that in many cases, all of the evidentiary material has been analyzed or consumed.

Final Comments - During the course of the discussion, the panel suggested the project staff inquire into:

(1) Examine trends of laboratory practice in terms of random re-analysis/re-testing and internal PT;
(2) Determine which types of agencies and persons submit cases to laboratories;
(3) Consider if all questions on the laboratory survey were answered with the same frequency by different types of laboratories (e.g., type of laboratory and if the laboratory was interested in participating in blind feasibility testing).
(4) The fraction of laboratories' personnel resources devoted to proficiency testing or other quality assurance procedures.
(5) Distinguish and weigh virtues of the different modes of PT and the roles each play in furthering QA/QC; and
(6) Develop protocol or manual for the review of blind PT results.

The meeting adjourned at approximately 1:00 pm.

C-3. SUMMARY OF THE THIRD NATIONAL FORENSIC DNA REVIEW PANEL (NFDRP) MEETING

Final Version (September 3, 1998)

June 22, 1998
New York, NY

NFDRP Members Present:

Ballantyne, J., Eisenberg, M., Gaertner, S., Hicks, J., Kuo, M., Lederberg, J., Murch, R., Reeder, D., Schanfield, M., Thompson, W., Weedn, V.
Rau, R., NIJ Project Monitor
Gaensslen, R., Peterson, J., and Lin, G., UIC Project Staff

NFDRP Members Absent:

Abrahamson, S., Budowle, B., Chakraborty, R., Devlin, B., Eisenberg, A., Ferrara, P., Laber, T., Scheck, B., Werrett, D.

The meeting was called to order by Dr. Lederberg.

Dr. Rau stated the charge to the panel and project staff for the meeting was to answer the following:

1) Is blind external proficiency testing feasible for DNA testing laboratories?
2) If yes, to recommend a national model external blind DNA proficiency testing program for the nation's public and private laboratories.

General

The meeting began with a general discussion of what is known about error rates in the forensic DNA testing area. Dr. Reeder noted that his review of declared proficiency studies reveal an approximate rate of error in the neighborhood of 1 in 1,000. The point was made that some have argued that if rates of error are greater than random match probabilities then the error rate needs to be expressed to the fact-finder in courts of law. If the rate of error is in the order of 1/1,000, then a great number of proficiency tests would be needed (more than 1 or 2 proficiency tests per examiner) to find such mistakes. The point was made by Dr. Thompson that error rates may vary greatly from laboratory to laboratory and from case to case. The question was raised if blind proficiency testing was the best means to identify/measure these variances. Also, the value of a "general rate of proficiency" was questioned if what is needed is a more focused measure, specific to the laboratory in question and sensitive enough to determine if the problem was analytical, clerical, or the result of sloppy evidence handling by police investigators or analysts.

Mr. Hicks noted that general validation data have been published for DNA techniques, and there was little evidence there is a serious problem in forensic DNA testing meriting the blind testing response. Dr. Thompson questioned the value of such validation data if they were generated from the examination of pristine samples, and did not reflect real world samples.

Dr. Lederberg raised the question if the DNA testing specialty was being unfairly singled out to supply such proficiency data while other forensic disciplines are not. It was noted that while presently courts of law do not require proficiency data to accompany other forms of forensic data, this may be changing as the result of decisions like Daubert v. Merrell Dow.

It was also noted that perhaps the best safeguard to protect against erroneous matching results being used in court (by prosecutors) would be the option of re-testing the sample by the defendant. If carried to the extreme, however, this might mean every sample would be tested twice and would have profound resource implications.

Dr. Gaensslen presented a summary of the draft final report that had been mailed to panel members prior to the meeting. His slide presentation covered:

1. The Definition of External Blind Proficiency Testing
2. Modes of Administration/Introduction of Blind PTs
3. The Status of Ten Blind Proficiency Tests Administered in this Project
4. Components of Laboratory Performance Assessable to Proficiency Testing
5. Acceptable Performance Criteria
6. Advantages and Disadvantages of Declared vs. Blind Proficiency Testing
7. Four Blind Test Introduction Modalities and Factors Defining Them as Internal/External
8. Difficulties with Tests at the Margin
9. Factors Affecting the Logistics of Large-Scale Blind Proficiency Testing Program
10. Cost Estimate Summary
11. Policy Recommendation Decision Flowchart
12. Future Research

Defining Feasibility and Blind Proficiency Testing
Given the charge to the panel, there was a discussion of the term "feasible" and its meaning in relation to such other terms as "possible" and "practicable." Webster's defines these terms as follows:

> Feasible – 1) capable of being carried out; 2) capable of being used or dealt with successfully; 3) reasonable, likely; synonym: see possible
> Possible – 1) Being within the limits of ability, capacity, or realization; 2) being what may be done or may occur;
> Practical – Capable of being put to use or account – useful; synonym: see practicable
> Practicable – possible to practice or perform, feasible; 2) capable of being used, usable

Possible implies that a thing may certainly exist or occur given the proper conditions; practicable implies that something may be easily or readily effected by available means or under current conditions; feasible applies to what is likely to work or be useful in attaining the end desired.

The panel members were generally in agreement that, for this project, "feasible" means more than simply "possible" and, as noted in the above definition, means "likely to work" or "be useful in attaining the end desired" but also encompasses such considerations as "cost" and the relative merits of blind testing vis-a-vis other quality assurance options.

The definition of blind proficiency testing expressed in the final report was:

1. A test presented to target laboratory through law enforcement/conduit laboratory in which the "case" was externally manufactured and no one in the target lab has advance information; OR
2. A test presented to DNA analysis unit in which "case" was externally manufactured and the fewest possible persons outside the DNA unit informed about the test; OR
3. A test by "random reanalysis" in which auditors from outside lab/system select case for reanalysis, audit/review all work done in case, and reanalyze evidence.

Relative Merits of Declared and Blind Tests

Dr. Lederberg asked if there was an "industry wide standard" as to the ability of "blind " testing to detect errors not discovered using more conventional "declared" proficiency testing. The project staff's review of the literature found blind proficiency studies generally reveal higher error rates than open (declared) testing, but the differences varied among the few studies comparing these differences. Studies found, for example, more false negative rates in urine drug screening in blind tests largely because less sensitive tests were used. No such studies are present in the DNA or forensic fields. Dr. Lederberg commented that, as an example, if blind testing detected errors at a rate 1,000 times greater than declared testing, then the field might be truly obligated to adopt that form of testing. But, given the values in the literature, he did not believe there was such a mandate. He also noted that at present the major argument for blind proficiency testing is less one that it is a superior measurement device and more one "to keep laboratory analysts on their toes" since analysts could never be sure the case before them wasn't a blind proficiency test.

There was also the question as to the frequency that the results of declared proficiency tests are used in court. There are no data available on this issue. Mr. Hicks asked how often DNA results constituted the only evidence against the defendant and expressed his belief that in most cases the DNA results were supported by other evidence.

Dr. Ballantyne noted that proficiency tests were being used as a regulatory device in New York State, and there was an instance in which a laboratory was temporarily closed down for poor proficiency test results.

Dr. Murch asked the question as to the relative merits of declared vs. blind testing. Dr. Weedn stated the systems are comparable except that blind testing might lead to 1) greater confidence in testing by outsiders, 2) heightened vigilance by laboratory personnel, and 3) an indication of how well the laboratories were performing the criminalistics/evidence selection aspects of case analysis.

Discussion also centered on the assumption that blind proficiency testing was superior because it "tests the whole system" better than declared tests, which merely test the analyst. Because blind testing presumably encompasses the handling, interpretation, and reporting of results it is viewed as a more comprehensive measure than declared proficiency testing. Still, it has limitations. A major limitation of blind testing is that it does not address the initial crime scene/evidence collection phase of the forensic testing process that is typically carried out by police investigators. This is because the blind proficiency materials are manufactured in a controlled environment and delivered to the law enforcement agency for packaging and integration of necessary evidence report forms. For this reason, a method like "random reanalysis", which entails the review of the entire completed case file (including reanalysis of the sample), actually allows a better, although not perfect, review of the initial phases of the process than blind testing. Even random or selected audits, which include a case file review, but not a reanalysis of the evidence, have been found to be extremely valuable.

Dr. Gaensslen discussed a slide citing two primary benefits of blind vs. declared testing: that blind testing is a more realistic measure than declared testing and that blind proficiency testing "tests the whole system." He noted that building "public confidence in laboratories" should be added to the list of attributes. Dr. Weedn noted that creating "heightened vigilance" on the part of examiners would be a fourth attribute. It was suggested that the presentation of the relative benefits of blind vs. declared testing be more clearly delineated in the report.

Ms. Gaertner noted that in her responsibilities as a prosecutor she has detected no lack of confidence by the general public in the work of the laboratories. If there are questions, it focuses at the initial stage of evidence gathering and the integrity of police investigators. Dr. Lederberg asked if there would be any test that would detect planted evidence and the consensus of the panel was there is none. Dr. Schanfield noted blind testing would not detect this type of problem. Dr. Reeder mentioned the possibility of searching for exogenous DNA on the evidence gathered by the police using highly sensitive techniques.

Dr. Murch questioned the ability of blind proficiency test providers to successfully simulate the truly problematic cases. Dr. Lederberg also questioned if the types of external factors/pressures present in some of these problematic cases can ever be replicated in blind proficiency cases.

Dr. Lederberg asked if the staff had investigated the various quality control measures used in the engineering field, and suggested the name of Warner North in Palo Alto.

Cost Considerations

On the discussion of cost estimates, several panel members expressed their belief that the estimated cost (included in the report) of $2,200 per blind proficiency test was unrealistically low and had excluded various administrative and logistical costs. For example, the travel costs were too low (instead of $500 should be closer to $1,500 per trip). Mr. Hicks noted that even in the "mature system" there may not necessarily be cost savings because over time the personnel in the local system might change, thereby necessitating the manufacturer to revisit the locale once again.

Under future research, Dr. Weedn suggested that UIC staff investigate if any manufacturer would be interested in engaging in such blind testing as a business practice. It may be that manufacturers would not be interested in such testing because the costs would be too great. He also suggested that staff attempt to validate the blind testing process by investigating how results would be used by participating laboratories to improve their operations.

Preservation of Evidentiary Samples for Reanalysis

Dr. Lederberg offered another legal option for framing DNA results, that being 1) to limit the use of DNA tests primarily as "investigative information;" and/or 2) to hold that DNA results would only be admissible in court if there was sufficient sample to permit defense testing of the evidence. What would be "lost" in terms of justice considerations under either or both of the above restrictions? Possibly by offering the defense the opportunity to re-analyze the evidence in every case could be equivalent to a "built in method of proficiency testing" that serves as a "random re-analysis" in the same sense as a "mechanical" system.

Whereas the legal process of "discovery" theoretically enables defendants to review the prosecution's scientific evidence and to secure that evidence for re-analysis, an obvious limiting condition is where the evidentiary sample is so small that it is consumed in the initial analysis. Here, the question was asked,

"How frequently is the entire biological sample consumed in the analysis by government laboratories and thereby not available for a second series of tests?" There are no data on this point, but this question might be investigated in the next phase of the present research.

Dr. Schanfield observed that in Colorado, in cases where there are no suspects, the state has no obligation to conserve a duplicate sample. He also noted that if the case doesn't appear on the court calendar, it will not be worked.

Dr. Ballantyne observed that, regardless of the size of the sample, there was no excuse not to split the sample in every case. The practice in his laboratory was always to preserve a portion (20%) of the sample in the event the defense wanted to perform retesting. The question was asked, however, if the splitting of very small samples might not actually lead to more questionable results, as the analytical systems are pushed to their limits and greater contamination possibly is introduced. Consequently, in cases where the evidentiary sample is very small, there is the question if it is better 1) to have one analysis by a government laboratory that consumes the entire sample, and produces a single data point; or 2) to attempt to divide the sample into two even smaller samples, and attempt to generate two data points produced by two different laboratories. The latter approach runs the risk of producing inferior or no results at all, because the divided samples are two small. These are questions that may merit additional research.

A discussion followed concerning the fraction of government scientific reports that are actually scrutinized by the defense. The belief this percentage is very small is an argument in favor of an initiative like blind proficiency test. Additionally, of those reports reviewed, what fraction has revealed problems, and what the major types of problems are. Dr. Thompson suggests the best way would be to query defense attorneys, academics, and scientists who are engaged in this type of defense work. In cases he has reviewed, he feels most problems concern interpretation issues, possible control failures, or where analysts do not interpret questionable results consistently from one case to the next.

Assessing the Feasibility of Blind Proficiency Testing

Discussion returned to the mandate of the panel, which is to recommend guidelines and that it is up to the DAB to issue standards. In terms of the feasibility of blind testing, when we consider the do-ability of blind testing, we confront a number of related issues that affect its do-ability. These range from those impinging on the ability to successfully introduce such tests into labs undetected (which was the focus of the present study), the length of time it takes laboratories to return results (in our project 3 of the 10 issued samples still have not been returned), to such issues as state laws that may forbid law enforcement officers from submitting false evidence, to whether a contractor would even bid on conducting such tests. In terms of the legality of submitting false evidence, Dr. Thompson felt that an overarching federal mandate for laboratories to engage in such testing would override state laws forbidding the manufacture of "false" evidence.

One of the challenges of introducing more complex blind proficiency samples, or so-called "tests at the margin," is the difficulty of producing multiple, identical samples. Even if the blind samples could be replicated at the manufacturing stage, with the passage of time, and as the evidence awaits processing, the samples would degrade to some extent (depending upon storage parameters) and would not remain identical indefinitely.

Continuing the discussion of problems associated with blind testing, the panel noted the problems inherent in the process including fabrication of the test samples, the introduction of fabricated evidence into databases, and then expunging it. Although every effort would be made to expunge traces of a blind

proficiency test from databases, it is inevitable, over time, that some samples would be missed and would remain in the system. Another problem is the substantial costs associated with blind testing. Dr. Reeder added that the experience of NIST is that it costs them about $7,000 per test, which includes about 100% overhead (actual tests being around $3,500). The CODIS implications mean we would require a number of donors to avoid cross hits. There are legal prohibitions in some states that technically forbid agencies from submitting "false" evidence. There is also the difficulty in creating cases that truly approximate the most problematic types of cases. Also, Dr. Lederberg expressed his aversion to a process that is at base deceitful, and requires officers not to be truthful with laboratories. There is also the ever present potential problem of police officers/ investigators "tipping off" analysts the case is actually a blind. At bottom then, as noted by Dr. Eisenberg, the panel must decide if the primary benefits of blind testing, which are to keep analysts on their toes and to build public confidence in the laboratory, override these problems. Ms. Gaertner raised the issue that engaging in such a system might actually lead the public to lose confidence in the laboratory because it was engaged in a process that relied on fabricated samples. Some members suggested they could foresee analysts' integrity being impugned as a result of their analyzing fabricated evidence at government expense.

Dr. Lederberg asked for a panel consensus on the feasibility of blind testing and all, except for Dr. Thompson, believed it was not feasible. Dr. Thompson summarized his arguments in favor of blind testing.

Random Reanalysis/Reaudit

A discussion of random reanalysis followed, and its relative merits compared with blind testing. The point was made that the selection of cases would not necessarily have to be random, but that selected cases could be chosen for review. Ms. Kuo indicated that the availability of a sample would be one criterion, but cases could also be chosen based on the contention the results were problematic. The legal status of the case, whether it was open, closed, or under appeal might also be criteria to consider. Cases where there were acquittals might also be selected. The possibility of ASCLD/LAB adopting audits/reanalyses as a criteria for accreditation was also discussed. There was discussion that a complete audit of the case might be equally or more valuable as an actual reanalysis of the evidence itself. Experts might be called in to review the case file carefully and could, if necessary, include retesting. Some members expressed the belief that most cases would not warrant an actual reanalysis.

The question was raised if a separate agency would need to be created to conduct these audits. Funds would be required and the present fees charged by ASCLD/LAB for accreditation would probably not be sufficient. Ms. Kuo noted that ASCLD was already swamped with work and probably could not take on this responsibility too. Dr. Eisenberg raised the issue if ASCLD had the standing of an impartial agency so that its work would be accepted by the defense bar and the court. Which agency would have the responsibility for compiling all the paperwork in a case during a random audit? Dr. Thompson saw such a responsibility as not different from responding to a discovery motion, but others thought it would be a significant burden. Ms. Kuo stated that the costs of such audits might be just as great as blind proficiency testing, with burdens falling on the laboratory in gathering the records and on a group of individuals who would review the case file. Dr. Lederberg felt that as a society we are more accustomed to such retrospective reviews/audits of records and would be more acceptable than the blind proficiency system that depends upon examiners being deceived by submitting clients.

The audits of case files that ASCLD currently performs are more procedural than what is being proposed which is more of a technical review of the file. Mr. Hicks raised the question if consideration of reanalyses/re-audits was within the purview of this panel since its primary charge was to investigate the

feasibility of blind proficiency testing. Dr. Rau replied that the panel did have this authority since it was charged with the added task of proposing model programs for the nation.

Dr. Lederberg suggested that if every analyst was audited on one of his/her cases per year, and if this audit required one person-day of effort per analyst, we might be looking at a cost of about 1-2% of the local laboratory's budget to support this. Dr. Thompson thought the random audit proposal, with the provision for re-analysis, might be a good technique for identifying those problematic cases. Dr. Thompson thought that he and others involved in reviewing DNA cases could be contacted to produce a list of cases that are problematic. He thought it essential that the reviews be external, and that results be publicly available (to build confidence in the process). Even if 1 of every 100 cases worked were audited, it could produce useful information. Even if cases were not found to have outright incorrect results, the detection of cases where sloppy work was done or where interpretations of data might be disputed would be useful to know.

Under random reanalysis there is the question as to where the evidence is stored after its initial analysis. In most situations, the evidence is returned to the submitting law enforcement agency that may store it in a variety of locations. The location of the evidence and the conditions under which it is stored are potential problems.

In any type of random reanalysis, standards/criteria must be developed that could be used in selecting cases for retesting and taking appropriate action thereafter.

Progress in Forensic DNA Testing

Dr. Lederberg also brought up the point that there have been substantial improvements in forensic DNA testing in the past several years, including the promulgation of standards and guidelines by TWGDAM, ASCLD, and the DAB, but that we don't really know yet if they will have the desired effect in the long term. He anticipated that a comprehensive accreditation program for forensic laboratories was on the horizon, as well. He suggested that perhaps we should wait for a period of time to allow these programs to work before introducing a requirement like blind external proficiency testing. Nonetheless, the field still needs a mechanism that can measure the success of these other programs, even if blind proficiency testing is not recommended.

Recommendation **

After reviewing the various options, Dr. Lederberg proposed the following resolution:

1. The accreditation system and associated quality assurance guidelines of the DNA Advisory Board needs to be given the opportunity to take hold

2. It is recommended that the DNA Advisory Board generate guidelines for more stringent external case audits for use by ASCLD-LAB, or another relevant accrediting body, as part of the accreditation process. The external case audits should be conducted regularly and serve as a measure of how well accreditation and its associated requirements are working in a quality assurance context.

3 In the extreme, blind proficiency testing is possible, but fraught with problems (including costs), and it is recommended that a blind proficiency testing program be deferred for now until it is more clear how implementation of the first two recommendations are serving the same purposes as blind proficiency testing.

** May still be subject to changes in language based on committee comment or review.

All panelists were in favor of the above statement except for Dr. Thompson who favored blind testing.

Dr. Lederberg also noted that the audits would encompass a wide range of individual programs underway and that the audits could not address the effectiveness of each of the programs individually. There are many other issues involved that will have to be addressed by the DAB on a continuing basis.

Dr. Gaensslen summarized several research issues that UIC may address in the coming 15-month project period.

1. Enhance laboratory participation rates in blind proficiency tests
2. Increase law enforcement agency participation
3. Prepare a set of more difficult/complex proficiency tests
4. Gather more data on random reanalysis and costs

In addition, the following ideas were presented as possible research directions for the UIC team:

Survey of Government DNA Laboratories

There was the suggestion that an additional survey of DNA testing laboratories could provide useful information. Possible lines of questioning included:

1. Do you believe blind proficiency test results would be helpful in improving your operations? Would they be more/less helpful than open/declared test results?
2. How often have you presented DNA proficiency test results of any type in court? Please describe this experience.
3. How often have DNA results served as the only evidence in a case?
4. How often do cases arise in which biological samples are so small that they are consumed in their entirety in DNA testing?
5. How often has your laboratory attempted to divide a very small sample, to enable duplicate defense testing, and this led to unsatisfactory laboratory results?
6. Do you, as a practice, always attempt to preserve a portion of the sample for possible re-analysis by the defense?
7. How often is there any judicial scrutiny of your DNA laboratory results by the defense? What is the nature of this review?
8. In cases where a defense expert has reviewed your DNA test results in what percentage was the evidence actually re-analyzed? Did these results agree/disagree with the original findings?
9. What is your policy with respect to the retention and storage of DNA evidence after analysis?
10. Are there any regulations/laws in your jurisdiction that would forbid law enforcement officers from submitting fictional cases in the form of blind proficiency tests to your laboratory for analysis?

Survey of Defense DNA Attorneys/Experts

1. In what percentage of DNA cases analyzed by government laboratories that you have reviewed do you believe the results and/or interpretations were problematic? What types of problems have you found? In what percent of these cases have you undertaken/arranged for an independent analysis?

Did these results agree/disagree with the state's results? What percent of cases should have had an independent re-analysis?

2. What has been your experience in attempting to secure funds (from the courts or other sources) to have DNA evidence re-examined?
3. In your experience, how often have you found that critical biological evidence was consumed in its entirety by the government laboratory in its initial DNA testing? Were you notified in advance/present during this examination? Were you satisfied with this process?
4. Of cases where you have had evidence re-examined, how often have the results been different from the initial prosecution's results?

Survey of Private Proficiency Test Manufacturers

Query (declared) proficiency testing manufacturers such as CTS, SERI, CAP, Cellmark, and ask if they would be interested in engaging in blind DNA testing. The suggestion was made to have them propose rough cost estimates.

UIC staff also needs to propose to DAB the criteria that need to be used in selecting cases for case audits, and for evaluating the cases themselves.

Dr. Lederberg tendered his resignation from the panel effective immediately.

C-4. SUMMARY OF THE FOURTH NATIONAL FORENSIC DNA REVIEW PANEL (NFDRP) MEETING

Madison Room, Double Tree Hotel
300 Army Navy Drive, Arlington, Virginia
November 16, 1999
6:00 pm to 9:00 pm

NFDRP Members Present:

Ballantyne, J.; Chakraborty, R.; Devlin, B.; Eisenberg, A.; Eisenberg, M.; Gaertner, S.; Hicks, J.; Kuo, M.; Schanfield, M.; Thompson, W.; and Weedn, V.
Rau, R., NIJ Project Monitor
Gaensslen, R.; Peterson, J.; Lin, G.; and Ho, M.: UIC Project Staff

NFDRP Members Absent:
Abrahamson, S.; Budowle, B.; Ferrara, P.; Laber, T.; Murch, R.; Reeder, D.; Scheck, B.; and Werrett, D.

The meeting was called to order by Dr. Peterson.

Dr. Peterson began the meeting by recounting the recent action of the National Institute of Health's Office for Protection form Research Risks (OPRR) and its impact on the University of Illinois at Chicago. On August 27, 1999, OPRR temporary suspended all federal research on the UIC campus.

The Blind Testing Project is largely unaffected since the bulk of data gathering and analysis had been completed by August 27th. Although any new research involving human subjects is suspended until the project protocol is re-reviewed and approved by the UIC Institutional Review Board, UIC's Office for the Protection of Research Subjects (OPRS) informed the proposal investigators that outstanding blind samples could remain in process since they had been disseminated prior to the date of suspension. However, re-approval must be acquired prior to publication of study results.

After informing panelists of the NIH decision, Dr. Peterson then introduced Dr. Rau of the National Institute of Justice (NIJ), who briefed the Panel on other on-going projects of NIJ and displayed recent publications.

Next Dr. Peterson outlined the purposes of the meeting as outlined in the agenda. The items of the agenda are:

- Review of Phase I recommendations that were decided during the previous NFDRP meeting (June 1998), as stated in the October 1998 Final Report to NIJ.
- Objectives of Phase II Research
- Discussion of Phase II Survey Results
- Status of Phase II Blind Proficiency Testing Results
- Discussion of Random Reanalysis
- Re-evaluation of Phase I Final Recommendations

Discussion of Phase I Recommendations

Dr. Peterson first directed the panelists to the Executive Summary in their information packet, and presented the three Phase I principal recommendations (as stated in the Executive Summary) for discussion:

Phase I Recommendations are:

1) The accreditation system and associated quality assurance guidelines of the DNA advisory board needs to be given the opportunity to take hold.

2) It is recommended that the DNA Advisory Board generate guidelines for more stringent external case audits for use by ASCLD-LAB, or another relevant accrediting body, as part of the accreditation process. The external case audits should be conducted regularly and serve as measure of how well accreditation and its associated requirements are working in a quality assurance context.

3) In the extreme, blind proficiency testing is possible, but fraught with problems (including costs), and it is recommended that a blind proficiency testing program be deferred for now until it is more clear how implementation of the first two recommendations are serving the same purposes as blind proficiency testing.

Dr. Rau noted that the Phase I Final Report had been forwarded to the Attorney General and a summary letter containing the three recommendations were sent to the U.S. House and Senate Judiciary Committees. Dr. Weedn expressed interest in receiving a copy of the said letter.

Objectives of Phase II Research

Dr. Peterson next informed panelists that objectives of Phase II of the Blind Proficiency Testing Feasibility Project, also outlined in the Executive Summary, are currently underway. The on-going tasks, including another laboratory survey, survey of defense attorneys, survey of expert witnesses, and another round of blind proficiency testing, will be discussed at the meeting today. These objectives were established to supplement Phase I findings by examining issues raised in the June 1998 NFDRP meeting. The five objectives of Phase II are as follows:

1) Examine current re-analysis programs of forensic DNA testing laboratories;

2) Perform additional field testing, using more complex scenarios;

3) Determine what fraction of worked DNA cases are reviewed and reanalyzed in forensic DNA testing laboratories;

4) Determine the extent to which original evidence items are still available for worked DNA cases that have been adjudicated; and

5) Explore the possibilities of a quantitative logistics analysis model to analyze blind proficiency testing model alternatives.

Discussion of Phase II Survey Results

I. DNA Laboratory Survey

Dr. Peterson presented a series of overhead transparencies describing the results of the Phase II laboratory survey. The survey was administered to 137 laboratories, and 91 were returned; of which, 67 laboratories reported to perform forensic DNA testing. Consequently, N=67 for the majority of tables in this section.

Panel members are asked not to circulate or reproduce the presented results until the findings have been re-checked for accuracy and finalized for dissemination.

1) Laboratories tested a median of 108 cases in calendar year 1997.

2) Laboratories were asked if, as a practice, they preserved a portion of the biological specimen or extracted DNA for subsequent confirmation or for retesting. Seventy percent of laboratories preserved both specimens and extracted DNA, 18% of labs preserve biological specimens only, 3% preserve extracted DNA only, and 9% do not preserve as a matter of practice.

   The panel suggested re-reviewing the questionnaires completed by laboratories that do not preserve specimens, for in cases where the evidence had to be presented in court, there would be a greater likelihood of preservation at the courthouse than cases which were dismissed before going to court. According to Dr. Schanfield, examination of courtroom retention policies might be better than looking at individual laboratories, for there is a wide range of policies.

   Panelists also commented on the finding that 70% of laboratories preserve both specimens and DNA extraction, and questioned whether laboratories answered the question prior to or after the implementation of DAB standards. Also, it was noted that the survey did not ask laboratories to specify the percentage portion of the specimen preserved in each case.

3) Laboratories were also asked the percentage of cases in which biological samples were so small that they were consumed in their entirety. About 73% of labs reported that between 0 and 10% of cases involved total consumption of biological samples. The median of cases including samples entirely consumed is 5%.

   Dr. Ballantyne instructed that the percentage of cases in which 0% of cases involved total consumption of samples should be specified. The panel also pointed out that it is important to note that the question did not distinguish between whether the entire sample in the case was consumed (and nothing was left) or if only one "sample" stain among many in a case was totally consumed.

4) We also asked after analysis if laboratories retained the original case evidence (e.g., the garment containing the stain) and/or a cutting containing the stain of interest. According to laboratories, 61% retains cuttings only, 2% retains the original evidence only, 10% retains both the original evidence and a cutting, and 27% retained none of the original evidence in the lab.

5) Given that a substantial portion of the evidence is returned to submitting agencies, it is not surprising that most (52%) labs don't know how long it will be retained. Those laboratories returning evidence to their own (agency) property rooms have a better idea of how long it will be retained.

6) We have gathered information on both internal and external QA/QC technical review and re-analysis activities. We initially asked the percentage of DNA analyzed cases in which there was an additional internal review of the case by peers, supervisors, or QA personnel. Ninety-nine percent of labs reported that between 99% and 100% of cases received such a review. The survey also asked the labs to estimate the percentage of cases receiving an external review and those in which defense counsel scrutinized laboratory data (lab notes, proficiency tests, etc.) beyond the lab report. In average, 18% of cases were subjected to defense scrutiny (median = 10%), while 12% of cases received external auditing and review (median = 10%).

   The panel noted that we do not know if the reported internal review was conducted by an analyst colleague prior to the writing of the final report or if it was a "past report" review. Furthermore, it is important to note that the mean number of cases analyzed by laboratories is 176, while the median number of analyzed cases is only 108. Therefore it might be possible that a few laboratories analyzing a great number of cases can inflate the average percentage of defense scrutiny and external review.

7) We also asked laboratories to estimate the percent of cases in which an analyst testified at a hearing or a trial. Laboratories reported an average of 15% of cases. There was considerable discussion at this point with some panelists thinking the rate should be higher and that it would vary by crime type and circumstances. Also, some questioned the difference between cases involving analyst testimony and cases involving defense scrutiny. UIC staff was instructed to compare the testimony estimate with the defense scrutiny estimate in an earlier question.

II. Defense Attorney Survey

Dr. Peterson next turned his attention to the survey of defense attorneys. Nineteen surveys were administered to defense attorneys with DNA litigation experience, but replies were received from only six. These six attorneys had served as the attorney of record in an average of 9 cases each that involved DNA evidence in calendar year 1997. As a group they had also served as "legal consultants" in a total of 20 additional cases.

Dr. Peterson first displayed the average percent of cases reviewed by attorneys by the types of reviews. Defense attorneys estimated that they reviewed test results in over 90% of cases, and that in more than 70% of cases they either consulted with an expert and/or had an expert review the test results. They reported that they visited laboratories in about two-thirds of cases, had DNA evidence re-tested in a quarter of cases, and tested additional samples in about 10% of cases.

Many panelists expressed their belief that this was not a typical sample of defense attorneys, for reported activities such as retesting of evidence and visiting laboratories are very uncommon in the real world. However, one panelist noted that defense attorneys often visit the laboratories. Nonetheless, panelists agreed that the results of the defense attorney survey represent the maximum possible values of the sample population. There was discussion if it would be worth attempting to increase the size of the sample. At minimum, we need to identify the six attorneys in our sample as a very select group. The small sample size was also of great concern to panelists.

Another transparency showed these defense attorneys estimated the original DNA lab results were problematic in 50% or less of cases and that they noted notable differences in 50% or more of the cases

they had re-analyzed. There was considerable discussion of these results and the belief by many panelists that these estimates of "problematic cases/results" were far too high.

III. Expert Witness Survey

Dr. Peterson then summarized the results of the expert witness survey. Twenty-seven experts were mailed questionnaires and eleven responded with data. These experts collectively reported reviewing 238 DNA cases for criminal defense lawyers in 1997, for an average of 22 cases per expert (median of 10). Two-thirds of the experts only consulted on their cases, and did not perform re-testing of evidence on cases reviewed. While majority of expert witnesses that performed re-testing found no significant difference between the original results and the re-test, majority reported that the total consumption of samples during original testing is a problem for re-testing. Furthermore, majority of experts reported finding "problematic results" and/or interpretations on some of the cases they reviewed in 1997.

Again, panelists pointed out that it is important to compare means and medians in the percentage of problems reported by expert witnesses, for the highly skewed number of reported cases by each expert can misrepresent the data.

Status of Phase II Blind Proficiency Testing Results

The podium was next turned to Dr. Gaensslen for discussion of Phase II Blind Proficiency Testing Results following the discussion on expert witness surveys. Dr. Gaensslen presented a status report on the 10 proficiency tests. According to Dr. Gaensslen, there is no one set way for laboratories to report DNA testing findings. Some laboratories report all information, while some laboratories only report findings. Of the tests returned so far, all have been accurate.

Discussion of Random Re-analysis and Case Re-analysis

Dr. Gaensslen reported that while random re-analysis is performed on only a small percentage of cases, all re-analysis result in identical findings as the original analysis. However, in the selection process of random re-analysis, cases in which evidence were totally consumed would be passed over. While random analysis is not really "random", according to Dr. Gaensslen, it is very extensive in documenting and recording the process. Dr. Gaensslen also added that the cost of random re-analysis is really not less costly than proficiency tests. It was suggested as a follow up on the subject that the investigators should visit a few laboratories and speak to the people who perform random reanalysis.

Dr. Weedn also pointed out that random re-analysis (CR) only works in large laboratories, while small laboratories would require personnel from other jurisdictions. The crossing of jurisdiction boundaries would turn CR into a contest, and there is no way to arbitrate the process because there is no right answer in these cases. Therefore he disagrees with mandating CR nationally. Furthermore, he argued that confirmations from random re-analysis are not as valid as proficiency tests, for it merely looks like two peers getting together and discussing the results.

Discussion of Final Recommendations

In reference to the Phase I recommendations, Dr. A. Eisenberg informed the Panel that the FBI DNA Advisory Board endorsed the first and the third recommendations, but not the second, for DAB does not want to guide accreditation. Dr. Kuo also added that ASCLB-LAB does not want to take on the policing role due to limited funding.

The Phase I second recommendation states:

It is recommended that the DNA Advisory Board generate guidelines for more stringent external case audits for use by ASCLD-LAB, or another relevant accrediting body, as part of the accreditation process. The external case audits should be conducted regularly and serve as measure of how well accreditation and its associated requirements are working in a quality assurance context.

Dr. Weedn pointed out that this recommendation is important in light of the congressional interest in blind proficiency testing because the forensic DNA testing community must review itself. According to Dr. Weedn, the purpose of this recommendation is to get labs to meaningfully and credibly look at each other and approve each other rather than mere self-assessment and reporting, just like any other professional/expert community which maintains its own standards.

The validity of the defense attorney and expert witness surveys was also greatly discussed, for panelists disagreed on how the findings should be presented in the final report, or if the findings should be reported at all. Some panelists argued that the findings from the defense attorney and expert witness surveys would confuse the community due to its weighted and limited sample. Problems that were brought up were the lack of internal system of validity and the lack of clear definition in self-reports. (For instance, what is the meaning of "alleged problems"?) Specifically, Dr. Devlin and The Honorable Ms. Gaertner felt that the two surveys offered no valuable information, for the opinions surveyed were that of public perception. However, Dr. Thompson countered by stating that the survey results were actual perceptions of the system by its participants. Dr. Schanfield suggested that the surveys should be described, but not used for its original purpose of evaluating levels of problems in forensic DNA testing laboratories.

The Panel instructed the investigators to return to the questionnaires and identify whether the "problems" identified by the defense attorneys and expert witnesses are substantive errors of the type that might be discovered by blind proficiency testing, and then cross-reference the alleged problems with the corresponding laboratories. Dr. M. Eisenberg wanted a description of the surveyed defense attorneys and expert witnesses, in terms of where they are in the field, their level of expertise in the community. Furthermore, Dr. Ballantyne stated that the critical question in surveying defense attorneys and expert witnesses is whether or not they found "false exclusion and/or false inclusion" in their case reviews.

Also, it was pointed out by the panel that the laboratory survey was conducted in 1997, possible before laboratories began compliance to DAB regulations, so that must be noted in the final report.

In the remaining time, Dr. Gaensslen pushed for a revision of the Phase I recommendations. Because DAB has already endorsed recommendations one and three, discussion centered on the re-wording of the second recommendation. It was noted earlier that DAB does not seek to recommend

guidelines to the community, and Dr. A. Eisenberg suggested using SWIGDAM instead of DAB or ASCLD-LAB. In general, panelists disagreed on whether or not to identify a specific agency (such as ASCLD-LAB) for the generation of guidelines, and the description of the guidelines.

After much discussion, the final recommendations were accepted to be as follows (changes are in CAPS):

1) The accreditation system and associated quality assurance guidelines of the DNA advisory board needs to be given the opportunity to take hold.

2) It is recommended that the SWGDAM generate guidelines for external case audits as part of the ANNUAL AUDITING process. The external case audits should be conducted regularly and serve as a measure of how well accreditation and its associated requirements are working in a quality assurance context.

3) In the extreme, blind proficiency testing is possible, but fraught with problems (including costs), and it is recommended that a blind proficiency testing program be deferred for now until it is more clear how implementation of the first two recommendations are serving the same purposes as blind proficiency testing.

Dr. Peterson thanked the panelists for their contributions to the project and adjourned the meeting.

D-1. PHASE 1 LABORATORY SURVEY INSTRUMENT - 1

University of Illinois at Chicago
Chicago, Illinois

Blind DNA Proficiency Testing Survey Instrument 1996

BASIC LABORATORY INFORMATION

In the space below, please enter the complete name and mailing address of your laboratory, and the name, telephone and FAX numbers, and email address of the person completing the form.

Name/address of laboratory _____

_____

_____

_____

Name of person completing survey: _____

Telephone ( )
FAX ( )
email

FOR THE FOLLOWING QUESTIONS, PLEASE CIRCLE THE NUMBER WHICH CORRESPONDS TO YOUR BEST RESPONSE. IF YOU CIRCLE OTHER PLEASE SPECIFY/EXPLAIN YOUR RESPONSE IN THE SPACE PROVIDED.

1. Does your laboratory perform only disputed paternity or disputed affiliation cases?

1 Yes 2 No (If the answer is No, go to Question 2.)

(If your answer is Yes, please do not answer any additional questions. Be sure your address and telephone/fax/email numbers are complete and return the questionnaire in the envelope. Thank you very much)

2. Does your laboratory perform DNA analyses on forensic evidence? (Please circle the number corresponding to your response.)

1 Yes (Please skip to Question 5.)
2 No
3 Not currently performing DNA analysis, but will be in the next 6-12 months.

3. If your laboratory does not perform DNA analyses, do you sometimes send cases or evidence out for DNA analysis?

1 Yes (Please go to Question 4.)
2 No

4. If yes, where do you send cases/evidence for DNA analysis? (Please circle all laboratories where you send evidence for DNA analysis, and return questionnaire in enclosed envelope.)

1 FBI Laboratory
2 Cellmark Diagnostics
3 Lifecodes Corporation
4 Roche Biomedical Labs
5 Forensic Science Associates
6 Serological Research Institute (SERI)
7 Other (please specify)

NOTE: THE FOLLOWING QUESTIONS SHOULD BE ANSWERED ONLY BY LABORATORIES THAT PERFORM THEIR OWN DNA ANALYSES ON FORENSIC EVIDENCE (BIOLOGICAL EVIDENCE FROM CRIMINAL CASES.)

5. Is your laboratory part of a larger state laboratory system?
1 Yes (If Yes, go to question 6.)
2 No (If No, skip to Question 8.)

6. Are there other laboratories in this system that provide DNA analyses?

1 Yes 2 No (If No, skip to 8.)

7. If Yes, will you be answering this questionnaire for:

1 Your laboratory only; or for
2 All laboratories performing DNA analysis in your system.

** If you circled (2) above, and are answering for all laboratories in your system performing DNA analysis, please do this consistently for all questions on this survey.

8. Excluding CODIS analysts, how many scientific personnel in your laboratory (or system)are engaged in DNA analyses (do not include supervisors, tech support staff, etc.)?

9. How many <u>total</u> scientific personnel are engaged in forensic testing of any type in this laboratory (or system)(do not include supervisors, tech support staff, etc.)?

10. How many of the analysts in this laboratory are American Board of Criminalistics(ABC) certified (any category)?

11. Is your laboratory within a (please circle the appropriate number)?

1 Local/County Police Dept
2 State Police/Highway Patrol
3 Dept Public Safety
4 Sheriff's Dept
5 Dept of Criminal Justice Services
6 District Attorney (or State Attorney)
7 Attorney General
8 Medical Examiner's Office
9 Dept of Health
10 Independent/Private Organization
11 Other (please identify)

12. If yours is a publicly funded laboratory, at what level of government is it located (please circle the appropriate number)?

1 Federal
2 State
3 County
4 City
5 Not applicable
6 Other (please describe)_____

13. Is your laboratory, or the larger laboratory system, American Society of Crime Laboratory Directors (ASCLD) accredited?

1 Yes 2 No (If No, skip to Question 16.)

14. If Yes, is your DNA section ASCLD accredited?

1 Yes 2 No 3 Not applicable

15. In what year was your laboratory (or system) first ASCLD accredited?

16. For approximately how many different agencies/clients do you provide forensic DNA services?

Approximate number of agencies/clients _____
Impossible to estimate _____

17. What is the approximate total population served by these agencies?

Approximate population _____
Impossible to estimate _____

18. Does this laboratory's jurisdiction have a standardized sexual assault evidence collection kit that is used in most cases for victims?

1 Yes (If Yes, go to Question 19.)
2 No (If No, skip to Question 20.)
3 Other (Please explain)

19. Who supplies the victim collection kit for use by police and emergency room personnel?

1 Crime laboratory
2 Police agency(ies)
3 Hospital
4 Victim services agency
5 Other (Please explain) _____

_____

20. Which sexual assault victim evidence collection kits do you regularly receive in your laboratory? Please identify (by brand name if possible).

21. Does this laboratory's jurisdiction have a standardized sexual assault evidence collection kit for suspects that is used most of the time?

1 Yes (If Yes, go to Question 22.)
2 No (If No, go to Question 23.)
3 Other (Please explain) _____
_____

22. Who supplies the suspect collection kit for use by police and emergency room personnel?

1 Crime laboratory
2 Police agency(ies)
3 Hospital
4 Victim services agency
5 Other _____
_____

23. In what form does this laboratory ordinarily receive specimens from suspects in sexual assault cases? Circle numbers corresponding to all answers that apply.

Blood
1 EDTA (purple top) tube
2 ACD tube
3 Clot tube
4 Dried bloodstain
5 Other (specify)_____

Saliva
1 Liquid
2 Dried stain on filter paper
3 Dried stain on cotton swatch
4 Oral swab
5 Other (specify)_____

Head and/or Pubic Hair Standards

In what type of container are standards submitted?

BIOLOGICAL EVIDENCE ACCEPTANCE POLICIES

24. Do you accept and attempt to DNA type ALL FORMS of biological evidence that may be potentially DNA typeable?

1 Yes 2 No (If Yes, please skip to Question 26.)

25. What types of specimens do you NOT accept?

1 Soft tissues
2 Bone
3 Teeth
4 Urine
5 Hairs (anagen)
6 Other (please identify)_____
_____
_____

26. Are there case circumstances under which you will NOT proceed with DNA typing on blood or physiological fluid evidence (e.g., if there is no suspect?)

1 Yes 2 No (If No, please skip to Question 28.)

27. This laboratory will NOT proceed with DNA typing on submitted specimens: Please circle all that apply.

1 In any type of case where there is no suspect.
2 In a blood comparison case where there is no suspect.
3 In a sexual assault case where there is no suspect.
4 In a case where questioned bloodstains were submitted without known bloods.
5 In sexual assault cases where vaginal swabs or semen are submitted without knowns from a suspected depositor.
6 In sexual assault cases if known blood from the victim was not submitted.
7 Other (Please explain)_____
_____
_____

INTAKE AND INITIAL PROCESSING OF DNA EVIDENCE

We wish to know how often various police and scientific personnel collect DNA typeable evidence from various locations for submission to your laboratory.

We want to know who actually collects and packages the DNA evidence you receive - not who delivers it to the laboratory. If you are a laboratory that ordinarily is not informed as to who actually collected the evidence, please skip Questions 28-33.

28. How often do the following personnel collect biological evidence from CRIME SCENES or from RELATED LOCATIONS pursuant to search warrants (e.g., seizure of suspect's clothing from his own house, etc.)

TYPE OF PERSONNEL	FREQUENCY THEY COLLECT BIOLOGICAL EVIDENCE (Please circle number corresponding to best response) Always / Most of the time / Occasionally / Never			
Patrol officers	1	2	3	4
Detectives	1	2	3	4
Uniformed officers or detectives with special training	1	2	3	4
Police evidence collection techs	1	2	3	4
Civilian evidence collection techs	1	2	3	4
Crime lab personnel	1	2	3	4
Other (Please specify)	1	2	3	4

_____

_____

29. How often do the following personnel collect biological evidence from VICTIMS AT HOSPITALS / MEDICAL FACILITIES?

TYPE OF PERSONNEL	FREQUENCY THEY COLLECT BIOLOGICAL EVIDENCE (Please circle number corresponding to best response) Always / Most of the time / Occasionally / Never			
Physicians	1	2	3	4
Nurses	1	2	3	4
Med techs	1	2	3	4
Physician's assistants	1	2	3	4
Other (Please specify)	1	2	3	4

_____

_____

30. How often do the following personnel collect biological evidence from CONSENTING SUSPECTS NOT UNDER ARREST? (If not applicable, skip and answer Question 31, "Yes")

TYPE OF PERSONNEL	FREQUENCY THEY COLLECT BIOLOGICAL EVIDENCE (Please circle number corresponding to best response) Always / Most of the time / Occasionally / Never			
Physicians	1	2	3	4
Nurses	1	2	3	4
Med techs	1	2	3	4
Physician's assistant	1	2	3	4
Evidence collection techs	1	2	3	4
Crime lab personnel	1	2	3	4
Other (Please specify)	1	2	3	4

_____

_____

31. Situation described in #30 not applicable/never occurs.

1 Yes 2 No

32. How often do the following personnel collect biological evidence from SUSPECTS UNDER ARREST (by warrant or court order).

TYPE OF PERSONNEL	FREQUENCY THEY COLLECT BIOLOGICAL EVIDENCE (Please circle number corresponding to best response) Always / Most of the time / Occasionally / Never			
Physicians	1	2	3	4
Nurses	1	2	3	4
Med techs	1	2	3	4
Physician's assistants	1	2	3	4
Evidence collection techs	1	2	3	4
Crime lab personnel	1	2	3	4
Other (Please specify)	1	2	3	4

_____

_____

33. How often do the following personnel collect biological evidence (actually collect the specimen from the bodies) from MEDICAL EXAMINER OR CORONER'S OFFICE?

TYPE OF PERSONNEL (Please circle number corresponding to best response Always / Most of the time / Occasionally / Never	FREQUENCY THEY COLLECT BIOLOGICAL EVIDENCE			
Autopsy pathologist	1	2	3	4
Autopsy technician	1	2	3	4
Coroner/coroner's asst.	1	2	3	4
Police officer/detective	1	2	3	4
Crime lab personnel	1	2	3	4
Other (Please specify)	1	2	3	4

_____

_____

RECEIPT OF BIOLOGICAL EVIDENCE BY THE FORENSIC LABORATORY

34. In this lab, biological evidence is usually taken in (accepted at the door) by:

1 Taken at the door by evidence clerk
2 Taken at the door by forensic examiner
3 Picked up by evidence clerk from police evidence custodian
4 Picked up by forensic examiner from police evidence custodian
5 From other forensic science laboratories (including in your own system)
6 Other (please specify)_____

_____

35. In this laboratory, biological evidence is received (circle all that apply):

1 Directly from an individual
2 U.S. Mail
3 UPS/FedEx/Other commercial carrier
4 Collected directly by a forensic examiner
5 Other (please specify)_____

_____

49. How important are the following factors in giving cases/specimens a higher priority (i.e. getting worked faster)?

FACTOR	IMPORTANCE (Please circle number corresponding to best response) Very Important	Somewhat Important	Unimportant
Date evidence received	1	2	3
Provide investigative leads	1	2	3
Needed for arrest warrant	1	2	3
Charging/prelim hearing deadline	1	2	3
Trial date	1	2	3
Seriousness of case	1	2	3
Willingness of prosecutor to use DNA typing results if provided	1	2	3
Other (please specify)	1	2	3

_____

_____

ASSIGNMENT OF CASES/SPECIMENS FOR ANALYSIS

50. How are cases/specimens assigned to particular DNA analysts?

1 Formal assignment to analyst by supervisor (Please answer Ques. 51.)
2 Informal rotation among analysts
3 Random assignment
4 Other (please describe)_____

_____

(Unless you circled "formal assignment", skip to Question 52.)

51. Who in this laboratory assigns cases/evidence to analysts?

1 Laboratory director
2 Deputy/asst lab director
3 Serology/biology/biochemistry unit supervisor
4 DNA unit supervisor
5 Other (please identify)_____

52. If a case has a number of different biological evidence items, are they all necessarily assigned to a particular DNA analyst, or could they go to more than one analyst?

1 All go to the same analyst
2 Could go to different analysts
3 Other (please specify)_____

_____

53. If one or more biological evidence items are subsampled, are the subsamples necessarily assigned to a particular DNA analyst, or could they go to more than one analyst?

1 All go to the same analyst
2 Could go to different analysts
3 Other (please specify)_____

_____

54. How is it decided which analysts will be assigned different evidence items (or subsamples from the same evidence item) from the same case?

1 Analyst availability (assignment to keep workloads even)
2 Based on specialization (analysts specialize in PCR, RFLP, etc)
3 Other (please specify)_____

_____

DNA ANALYSIS OF CRIMINAL-CASE EVIDENCE

55. Does this laboratory do conventional serological testing on biological evidence?

1 Yes 2 No (If No, skip to Question 57.)

56. Is conventional serological testing (e.g., ABO, isoenzyme, serum protein typing) done before DNA testing is considered or initiated?

1 Yes 2 No 3 Sometimes

If No or Sometimes, please explain. _____

_____

57. Does this laboratory's DNA testing consist of:

1 RFLP only (Go to Question 58.)
2 PCR-based testing only (Go to Question 66.)
3 Both RFLP and PCR-based testing (Go to Question 68.)

QUESTIONS 58-65 ARE FOR LABS THAT DO RFLP TESTING ONLY

58. Do you follow the "FBI Methods" for RFLP or modifications of them?

1 Yes 2 No

59. How many loci do you have the capability of typing?

60. How many loci do you generally type?

61. Which loci are you capable of typing? (circle all that apply)

1 D1S7
2 D2S44
3 D4S139
4 D5S110
5 D10S28
6 D17S79
7 D17S26
8 Other (Please specify)

62. Which loci do you generally type? (circle all that apply)

1 D1S7
2 D2S44
3 D4S139
4 D5S110
5 D10S28
6 D17S79
7 D17S26
8 Other (Please specify)

63. Does this lab use chemiluminescence detection?

1 Yes 2 No

64. If the DNA is of insufficient quantity or quality for RFLP, do you send specimens elsewhere for PCR-based typing?

1 Yes (If Yes, go to Question 65.)
2 No (If No, go to Question 78.)

65. Where do you send the specimens for PCR-based typing?

QUESTIONS 66-67 ARE FOR LABORATORIES THAT DO PCR-BASED TESTING ONLY.

66. What are your capabilities?
(If you circle 7 "STRs", please answer Question 67. If not, skip to Question 78.)

1 HLA-DQA1
2 AmpliType® PM (LDLR, D7S8, GYPA, HBGG, GC)
3 D1S80
4 XY (alphoid centromeric repeat sequence)
5 XY (amelogenin)
6 ZFX/ZFY
7 STRs (Please answer Question 67.)
8 Other (Please specify)

67. Which STR loci do you have the capability of typing?

1 CSF1PO
2 TPOX
3 THO1
4 HPRTB
5 vWF
6 FESFPS
7 F13A01
8 F13B
9 LPL
10 D21S11
11 D18S51
12 D8S1179
13 Other (specify)

QUESTIONS 68-77 ARE FOR LABS THAT DO BOTH RFLP AND PCR-BASED DNA TYPING.

68. Do you use PCR-based typing to screen evidence before deciding whether to do RFLP or additional PCR-based typing?

1 Yes, always (Go to Question 69.)
2 No, never (Skip to Question 70.)
3 Sometimes, other criteria involved in deciding (Skip to Question 70.)

69. Which loci are used initially, for the screening?

1 HLA-DQA1
2 AmpliType® PM
3 D1S80
4 XY (gender) sequences
5 STRs
6 ABO genotyping
7 Other (specify)

70. Do you follow the "FBI Methods" for RFLP or modifications of them?

1 Yes 2 No

71. How many loci do you have the capability of typing?

72. How many loci do you generally type?

73. Which loci are you capable of typing? (circle all that apply)

1 D1S7
2 D2S44
3 D4S139
4 D5S110
5 D10S28
6 D17S79
7 D17S26
8 Other (Please specify)

74. Which loci do you generally type? (circle all that apply)

1 D1S7
2 D2S44
3 D4S139
4 D5S110
5 D10S28
6 D17S79
7 D17S26
8 Other (Please specify)

75. Does this lab use chemiluminescence detection?

1 Yes 2 No

76. What are your PCR-based typing capabilities? (If you circle 7 "STRs" answer Question 77. If not, go to Question 78.)

1 HLA-DQA1
2 AmpliType® PM (LDLR, D7S8, GYPA, HBGG, GC)
3 D1S80
4 XY (alphoid centromeric repeat sequence)
5 XY (amelogenin)
6 ZFX/ZFY
7 STRs (Answer Question 77)
8 Other (Please specify)

77. Which STR loci do you have the capability of typing?

1 CSF1PO
2 TPOX
3 THO1
4 HPRTB
5 vWF
6 FESFPS
7 F13A01
8 D21S11
9 LPL
10 D21S11
11 D18S51
12 D8S1179
13 Other (specify)

LABORATORY NOTES

78. Do DNA analysts in this lab keep laboratory notes (notes that are kept in the lab, and used to formulate the Lab Report)?

1 Yes 2 No (If No, go to Question 81.)

79. Are the internal laboratory notes:

1 Free form (analysts write notes on blank pages)
2 Fill-in type forms (forms exist, analysts fill in data)
3 Combination of free form and fill-in type forms
4 Other (specify)_____

_____

IF YOU USE FORMS, PLEASE PROVIDE A COPY.

REPEAT TESTING

80. Are the lab notes reviewed by another analyst or supervisor?

1 Yes 2 No

81. In this lab, are specimens re-tested to confirm results (assumingsufficient DNA?)

1 Yes 2 No (If No, skip to Question 83.)

82. Are the specimens re-tested by a different, or the same, examiner who did the original typing?

1 Same examiner explicitly
2 Different examiner explicitly
3 Could be either the original examiner or a different examiner

83. Are case specimens that were previously worked sometimes given to analysts to re-test as a QA/QC measure?

1 Yes 2 No

84. How frequently are case specimens that this lab has tested re-tested by another laboratory (assuming sufficient DNA)?

1 Always
2 Often
3 Occasionally
4 Never (go to Question 86.)

85. Has there ever been a typing discrepancy on the same specimen between your results in a case and results obtained by another laboratory?

1 Yes 2 No 3 Don't know

86. Does this laboratory routinely save a sample of each specimen (assuming there is enough) for possible future re-testing?

1 Yes 2 No (If No, skip to Question 88.)

87. Do you save for possible future analysis:

1 Biological evidence specimen itself if not consumed
2 DNA extracted from the biological evidence specimen if not consumed
3 Both (1 and 2 above)
4 Other (please explain)_____
_____

LABORATORY REPORTS (the report that is sent out to the submitter, the prosecutor, etc.)

88. We assume that the Lab Report routinely lists the date, case number(s), item number(s) and possibly brief description(s) of the evidence item(s), and the name(s) of the analyst(s). Besides that information, what results does this lab report? (Please circle number of responses that apply.)

1 Report RFLP band sizes for all specimens tested
2 Report RFLP band sizes for evidential specimens, but not known specimens
3 Report RFLP band sizes for known specimens, but not evidential specimens
4 Do not report RFLP band sizes at all
5 Report PCR-based test types (HLA-DQA1, PM, etc.) for all specimens tested
6 Report PCR-based test types for evidential specimens, but not known specimens
7 Report PCR-based test types for known specimens, but not evidential specimens
8 Do not report PCR-based test types at all

89. Do you report conclusions based on the typing results (which evidential specimens match/are the same type as which knowns, etc)?

1 Yes 2 No

90. Do you report the frequencies of the types found in the evidential specimens and/or in the knowns?

1 Yes 2 No (If No, skip to question 92.)

91. Do you report the frequencies in several ethnically distinct population groups (e.g. White, Black, etc)?

1 Yes 2 No

92. Does a laboratory report have to be reviewed by another analyst or a supervisor before it can be issued?

1 Yes 2 No

93. Does a laboratory report have to be approved by a supervisor before it can be sent out?

1 Yes 2 No

94. To whom is the laboratory report automatically sent?

1 Not automatically sent to anyone
2 Submitting agency or agent
3 Prosecutor (or defense counsel, if defense case)
4 Other (specify)_____
_____

95. Do analysts communicate with police officers, prosecutors (or defense counsel, if defense case) informally about cases that are in the laboratory?

1 Always
2 Frequently
3 Sometimes
4 Never

96. Will you give police/prosecutors (or defense counsel, if defense case) results over the phone once the case has been worked, but before the report is issued?

1 Yes 2 No (If No, skip to Question 98.)

97. The DNA test results are provided to police/prosecutor (or defense counsel, if defense case) over the phone:

1 Automatically in every case
2 We call them if it's an important case, but not routinely
3 They usually have to call and ask
4 Other(please explain)_____

_____

DATABANKING/CODIS

98. Is this laboratory involved in processing specimens for a state statute-mandated databank? (If No, skip to Question 124.)

1 Yes 2 No 3 Other (Please explain)

If yes, would you provide a copy of the statute (or the appropriate citation)?

99. Is this laboratory actually typing these specimens?

1 Yes
2 No, currently preparing and storing but not typing databank specimens (Please skip to Question 113.)
3 Other (Please explain)_____

_____

100. The databanking statute covers offenders convicted of:

1 Any felony
2 Serious felonies
3 Felonies against persons
4 Felonious sexual assault - any degree
5 All sexual assaults - felony or misdemeanor
6 Other (specify)_____

_____

101. The state's database is maintained by:

1 This laboratory
2 Other laboratory that is part of our system
3 Another laboratory not part of our system
4 A contractor (specify)_____
5 Other (specify)_____

102. This laboratory enters data into the database:

1 Directly, through a PC or terminal connection
2 Has to be sent to the database central location and entered there
3 Other (specify)_____
_____

103. Who is authorized to enter data into the database?

1 Any analyst
2 Only a supervisor
3 Only a specified individual
4 Not applicable - data is not entered from this laboratory
5 Other (identify)_____

104. Searching the database for a profile:

1 Can be done here on a PC or terminal connection
2 Has go to be done at the database central location
3 Other (specify)_____
_____

105. Who is authorized to search the database?

1 Any analyst
2 Only a supervisor
3 Only a specified individual
4 Not applicable - searching is not done here
5 Other (identify)_____

106. Do the same analysts who do criminal casework also do databank specimen typing?

1 Same analysts do both all the time (Skip to question 108)
2 Analysts rotate between casework and CODIS
3 This lab has separate, dedicated CODIS analysts
(If this is your answer, please answer questions 107 and 108.)
4 Other (Please explain)_____

_____

107. How many dedicated CODIS analysts do you have?

108. How many of your CODIS analysts are American Board of Criminalistics (ABC) certified (in any category)?

109. How many RFLP loci are databased? (If none, skip to Question 111.)

110. Which RFLP loci are databased?

1 D1S7
2 D2S44
3 D4S139
4 D5S110
5 D10S28
6 D17S79
7 D17S26
8 Other (Please specify)

111. Which PCR-based loci are databased? (If none, skip to Question 113.)

1 HLA-DQA1
2 AmpliType® PM (LDLR, D7S8, GYPA, HBGG, GC)
3 D1S80
4 XY (alphoid centromeric repeat sequence)
5 XY (amelogenin)
6 ZFX/ZFY
7 STRs (Please answer Question 112.)
8 Other (Please specify)

112. Which STR loci are databased?

1 CSF1PO
2 TPOX
3 THO1
4 HPRTB
5 vWF
6 FESFPS
7 F13A01
8 D21S11
9 LPL
10 D21S11
11 D18S51
12 D8S1179
13 Other (specify)

113. What specimens are collected from convicted persons for CODIS databanking?

1 Blood preserved in EDTA (purple top) tube
2 Blood preserved in ACD tube
3 Blood in clot tube
4 Buccal swabbing
5 Dried bloodstain
6 Other (specify)_____

114. Who collects databank specimens?

1 Law enforcement personnel
2 Department of corrections personnel
3 Forensic lab personnel
4 Other (specify)_____

115. Is a written receipt provided by this laboratory for CODIS specimens?

1 Yes 2 No

116. Are there special forms for accepting CODIS specimens into the lab?

1 Yes 2 No

If Yes, we would appreciate having a copy of these forms.

117. How are CODIS specimens assigned to a particular analyst?

1 No formal assignment - analysts take them on a rotating basis (to keep a balanced workload)
2 Assigned by a supervisor
3 Other (specify)_____

_____

118. This laboratory has:

1 No backlog of CODIS specimens (Skip to Question 120.)
2 A backlog of CODIS specimens

119. How are CODIS specimens prioritized for typing?

1 First in, first out
2 By release/probation/parole dates of the offenders to be typed
3 By seriousness of the offense committed
4 Other (specify)_____

_____

120. Are CODIS specimen typing results reviewed by a supervisor or another scientist before entry into the database?

1 Yes 2 No

121. Is CODIS specimen typing repeated as a QA/QC measure before entry into the database?

1 Yes, every specimen
2 Yes, a randomly selected percent of specimens
3 No (If No, skip to 123.)

122. Is the repeat typing performed:

1 Explicitly by the same examiner who did the first typing
2 Explicitly by a different examiner
3 Could be either the original examiner or a different one
4 Other

123. Approximately how many profiles are in your state's database?

YOUR LABORATORY'S QA/QC PROGRAM AND PROFICIENCY TESTING (PT)
(Pertains to all DNA typing sections of laboratory.)

124. Does this laboratory:

1 Follow TWGDAM QA/QC Guidelines
2 Have its own QA/QC program
3 Other

If YOUR LABORATORY HAS A WRITTEN POLICY DESCRIBING YOUR DNA QA/QC PROGRAM, WOULD YOU PLEASE SEND US A COPY?

INTERNAL PROFICIENCY TESTING (PT) (QUESTIONS 125-129)

125. Casework examiners in this lab do internal PTs?

1 Yes 2 No (If No, skip to question 130.)

126. How many internal PTs does each examiner complete per year?

127. Do you:

1 Manufacture in-house an internal PT
2 Use specimens from a previously analyzed case
3 Other (please specify)_____

_____

128. When PT results are reviewed with the examiner by the supervisor or QA/QC coordinator:

1 Only the DNA testing results are discussed
2 DNA test results as well as evidence receipt and handling, and reporting of results are discussed
3 Some other combination (specify)_____

_____

129. Is there separate internal PT program for CODIS analysts?
1 Not applicable, we have no CODIS analysts
2 We have a CODIS program and CODIS analysts participate in the same internal PT program as casework scientists
3 We have a CODIS program and CODIS analysts participate in their own internal PT program
4 Other (Please explain)_____

_____

EXTERNAL OPEN (DECLARED) PT PROGRAM (QUESTIONS 130-133)
(By "OPEN," we mean the examiner knows he/she is being tested.)

130. Do casework examiners in this lab do external open PTs?

1 Yes 2 No (If No, skip to question 142.)

131. How often is each casework examiner tested per year?

132. External PTs for this laboratory are supplied by:

1 Collaborative Testing Services (CTS)
2 College of American Pathologists (CAP)
3 Cellmark Diagnostics (IQAS)
4 Serological Research Institute (SERI)
5 Other (specify)_____

133. Is there a separate external PT program for CODIS analysts?

1 This lab does not have separate CODIS analysts
2 This lab has CODIS analysts but no separate PT program for them
3 This lab has a separate program for CODIS analysts
4 Other (specify)_____

EXTERNAL BLIND PT PROGRAM (QUESTIONS 143-141)
(By "BLIND," we mean the examiner does not know he/she is being tested.)

134. Do casework examiners in this lab perform external blind PTs?

1 Yes 2 No (If No, skip to question 142.)

135. Is blind PT done as part of your QA/QC program:

1 Regularly
2 Occasionally (skip to Question 137.)
3 Never (skip to Question 139.)

136. If regular, how often is each examiner blind externally proficiency tested per year?

137. If occasional, how many times have you done blind proficiency tests?

138. External, blind PTs are manufactured for this laboratory by:

1 CTS
2 CAP
3 Cellmark
4 SERI
5 Another forensic laboratory
6 Other (specify)_____

139. Is there a separate blind external PT program for CODIS analysts?

1 Not applicable, this lab does not have separate CODIS analysts
2 This lab has CODIS analysts but no separate PT program for them
3 This lab has a separate program for CODIS analysts

140. Your assessment of your experience with external blind PT:

1 This laboratory has not done it at all
2 It was successful in the sense that the tested examiner did not figure out that it was a PT
3 It failed in the sense that the tested examiner figured out that it was a PT
4 Other (Please explain)_____
_____

141. Did supervisors inform the analyst who was blind tested that he/she had been tested and what the results were?

1 Yes
2 No

142. We would appreciate receiving any other comments, suggestions or questions you may have regarding proficiency testing in general or blind proficiency testing of DNA analysts in particular.

143. If your laboratory was to participate in a fully blind DNA proficiency test and the fictitious case went through the laboratory undetected, what paper documents and/or electronic files would be created (and thus potentially have to be purged)? Do you believe this would be a problem?

144. Attached to this questionnaire is a listing of other DNA typing laboratories that we know about in your state. Would you mind reviewing this list and add the names of any other laboratories performing forensic DNA testing in your jurisdiction, or any other necessary corrections?

145. Would this laboratory be interested in participating in a small-scale, actual, blind proficiency testing trial program?

Participation means that your laboratory will be placed on a list of laboratories which might: (a)receive a "fictitious" case disguised to look real; (b) be asked to serve as a reference laboratory in a proficiency test; or (c) both.

1 Yes 2 No

If you answered Yes, please read and complete the enclosed Agreement (colored paper). The Agreement is separate and distinct from this questionnaire. We realize that administrative review may be required before a final decision can be made on whether this lab will participate or not. However, please return the survey instrument as soon as possible, and indicate to us if the Agreement form is under review.

146. Thank you very much for completing this questionnaire. Once we have received and tabulated results, we will be pleased to send you a copy of our findings. Please indicate below if you would like to receive a copy.

1 Yes, please send us a copy of results
2 No, we are not interested

D-2. PHASE 1 LABORATORY SURVEY INSTRUMENT - 2

July 16, 1997

«Contact_Person»
«OrganizationName»
«Address»
«Address1»
«City», «State» «PostalCode»

Dear :

As you know, we are conducting a research study, funded by the National Institute of Justice (NIJ), to explore the feasibility and practicality of nationwide blind DNA proficiency testing. In December 1996 we attempted to survey all forensic DNA testing laboratories across the country, including yours. From that survey, we identified approximately 40 laboratories that would be willing to be blind tested in our study. We then asked those laboratories to supply us with contact persons and addresses of the primary agencies that submit biological materials to those laboratories for DNA typing. The «Testing_Laboratory» has indicated it would be willing to be part of the present study and to be "blind" tested. They also informed us that your laboratory routinely submitted biological materials to them for DNA typing. The purpose of this letter is (1) to determine if your laboratory still submits evidence to the «Testing_Laboratory» and, if so, the types of cases and/or situations when you do; and (2) to explore your willingness to serve as a supplier of blind samples to this laboratory in the future.

For a blind testing program to work, we believe it would be necessary for participating second agencies/laboratories to cooperate in helping introduce specimens into trial-site laboratories in the guise of real cases, complete with all the appropriate and expected evidence containers, numbers, forms, and related paperwork. To do this will clearly require the cooperation of second agencies that normally submit physical evidence to that laboratory. In one model, we would introduce evidence disguised as a real case through a police agency to the laboratory. In another, we would introduce evidence disguised as a case through a "conduit" laboratory (i.e., a laboratory that either does not perform DNA typing or does not perform a form of DNA typing they desire to be completed.

The participating conduit laboratory would require that designated personnel advise the proficiency test preparers in detail about the usual procedures for biological-evidence case submission. For the "case" to be seen as routine by the lab, it would have to look routine in every respect. In addition, target laboratory personnel could question participating conduit laboratory personnel about the "case," and satisfactory responses would be necessary to avoid suspicion on the part of lab personnel.

We don't foresee participating conduit laboratories having to make any large expenditure of personnel, time or funds. Participation would typically require a meeting at which the proficiency test preparer could receive a detailed briefing on the type of case most likely to satisfy the requirements and criteria noted above. At the conclusion of the meeting, the preparer would have sufficient information to be able to manufacture the necessary biological evidence to support the hypothetical "case." At a future time, detailed arrangements would be made with the conduit laboratory to provide the designee with the "evidence" for submission to the laboratory in the standard way. Conduit laboratories would be reimbursed as necessary for any reasonable costs incurred as a result of participating in such a program.

It is highly likely that the national DNA Advisory Board will utilize our findings from this study to decide whether to require all forensic DNA testing laboratories to incorporate blind proficiency testing into their quality-assurance protocols. Accordingly, it is an important project, and its findings will make a significant contribution to the future shape of forensic DNA laboratory quality-assurance policies.

We would like to know if your laboratory would agree to assist in submitting a manufactured "case" and "evidence" to the forensic science laboratory to which you normally submit evidence, provided you knew that the lab was either agreeable or required to be blind tested.

We are not asking in this survey if your laboratory wants to become an actual participant in this study at this time. Our purpose is to find out whether your agency would agree in principle to participate in blind proficiency testing, or be opposed to it. We have enclosed a survey sheet and a self addressed, stamped envelope. If you could send it back to us at your earliest convenience, we would be very appreciative. We will tabulate the aggregate results from the agencies we have contacted and include this information in the final report. We will not identify you or your agency. If you have any questions or comments about the project, please feel free to call either of us.

Cordially,

Joseph L. Peterson, D.Crim. (312)413-0439
R.E. Gaensslen, Ph.D. (312)996-2250
Project Directors

Enclosure

NAME OF PERSON COMPLETING SURVEY: _____

TELEPHONE #:_____

NAME OF YOUR LABORATORY: _____

1. Does your laboratory routinely submit biological evidence to another laboratory for DNA testing? Please circle "Yes" or "No."

   Yes No

   If your answer is "No", please do not answer any additional questions. Return the questionnaire in the envelope. Thank you very much.

2. If your answer is "Yes", please tell us the types of cases and/or situations in which you do.

3. Would you agree to cooperate with proficiency test manufacturers in submitting a manufactured "case" and "evidence" as a blind proficiency test to the forensic science laboratory in your jurisdiction, provided the laboratory was agreeable or required to participate in such a program? Please circle "Yes" or "No."

   Yes No

4. If you answered "No", please tell us your reasons. If you have any additional comments or concerns we would like to know them as well.

D-3. PHASE 1 LAW ENFORCEMENT AGENCY SURVEY INSTRUMENT

«Contact_Person»
«OrganizationName»
«Address»
«Address1»
«City», «State» «PostalCode» July 22, 1997

Dear :

We are conducting a research study, funded by the National Institute of Justice (NIJ), to explore the feasibility and practicality of nationwide blind DNA proficiency testing. This project is a response to a Congressional mandate to the NIJ, part of the DNA Identification Act of 1994, to investigate blind external (DNA) proficiency testing for the nation's forensic DNA typing laboratories. As part of the research, we have surveyed the forensic DNA laboratories across the country and the «Testing_Laboratory» had indicated it would be willing to be part of the present study and to be "blind" tested. They also informed us that your agency routinely submitted biological materials for DNA typing. The purpose of this letter is to explore the willingness of your agency to be involved in such blind testing in the future.

The DNA Identification Act of 1994, that was part of the so-called "Crime Bill" (P.L. 103-322) requires, among many other things, the NIJ to explore the feasibility of a national program of blind proficiency testing for forensic science laboratories engaged in DNA typing. The law defined a "blind" proficiency test as one "..... presented to a forensic laboratory through a second agency and appears to the analysts to involve routine evidence." The impetus underlying this legislation is the feeling in some quarters that "blind" proficiency testing provides a better measure of a laboratory's quality-assurance procedures and the accuracy of its results that "open" or "declared" proficiency testing in which that lab knows that the specimens are a test. Open proficiency testing is now very common in the nation's forensic science labs, and is required for accreditation of a laboratory by the American Society of Crime Laboratory Directors and for certification of individual examiners by the American Board of Criminalistics.

For a blind testing program to work, we believe it would be necessary for law enforcement agencies to cooperate in helping introduce specimens into trial-site laboratories in the guise of real cases, complete with all the appropriate and expected evidence containers, numbers, forms, and related paperwork. This would clearly require the cooperation of a law enforcement agency that normally submits physical evidence to the target laboratory. Law enforcement agency participation would require that designated personnel advise the proficiency test preparers in detail about the usual procedures for biological-evidence case submission. For the "case" to be seen as routine by the lab, it would have to look routine in every respect. In addition, laboratory personnel could question participating detectives about the "case," and satisfactory responses would be necessary to avoid suspicion on the part of lab personnel.

We don't foresee participating law enforcement agencies having to make any large expenditure of personnel time or funds. Participation would typically require a meeting at which the proficiency test preparers could receive a detailed briefing on the type of case most likely to satisfy the requirements and criteria noted above. At the conclusion of the meeting, the preparers would have sufficient information to be able to manufacture the necessary biological evidence to support the hypothetical "case." At a future time, detailed arrangements would be made with the law enforcement agency to provide your designee with the "evidence" for submission to the laboratory in the standard way. Law enforcement agencies would be reimbursed as necessary for any reasonable costs incurred as a result of participating in such a program.

It is highly likely that the national DNA Advisory Board will utilize our findings from this study to decide whether to require all forensic DNA testing laboratories to incorporate blind proficiency testing into their quality-assurance protocols. Accordingly, it is an important project, and its findings will make a significant contribution to the future shape of forensic DNA laboratory quality-assurance policies.

We would like to know if your agency would agree to assist in submitting a manufactured "case" and "evidence" to the forensic science laboratory to which you normally submit evidence provided you knew that the lab was either agreeable or required to be blind tested.

We are not asking in this survey if your laboratory wants to become an actual participant in this study at this time. Our purpose is to find out whether your agency would agree in principle to participate in blind proficiency testing, or be opposed to it. We have enclosed a survey sheet and a self addressed, stamped envelope. If you could send it back to us at your earliest convenience, we would be very appreciative. We will tabulate the aggregate results from the agencies we have contacted and include this information in the final report. In this summary, we will not identify you or your agency. If you have any questions or comments about the project, please feel free to call either of us.

Cordially,

Joseph L. Peterson, D.Crim. (312) 413-0439
R.E. Gaensslen, Ph.D. (312) 996-2250
Project Directors

Enclosure

NAME OF PERSON COMPLETING SURVEY: _____

TELEPHONE #: _____

NAME OF YOUR AGENCY: _____

Would you agree to cooperate with proficiency test manufacturers in submitting a manufactured "case" and "evidence" as a blind proficiency test to the forensic science laboratory in your jurisdiction, provided the laboratory was agreeable or required to participate in such a program? Please circle "Yes" or "No."

Yes No

If you answered "No", please tell us your reasons. If you have any additional comments or concerns we would like to know them as well.

D-4. PHASE 2 LABORATORY SURVEY INSTRUMENT

University of Illinois at Chicago
Forensic Science Program
Chicago, Illinois

DNA LABORATORY SURVEY INSTRUMENT 1998

GENERAL LABORATORY INFORMATION

In the space below, please enter the complete name and mailing address of your laboratory, and the name, telephone and FAX numbers, and e-mail address of the person(s) completing the form.

Name/address of laboratory: _____

_____

_____

_____

Name of person(s) completing survey: _____

Telephone: (___)_____
FAX: (___)_____
e-mail: _____

For the following questions, please **circle the number** which corresponds to your best reponse. If you circled OTHER, please specify/explain your response in the space provided.

1. Does your laboratory currently perform DNA analyses on forensic evidence?

a) Yes
b) No

If your answer is No, please do not answer any additional questions. Be sure your address and telephone/fax/email numbers are complete and return the questionnaire in the envelope. Thank you very much.

2. If your laboratory is part of a larger state laboratory system and other laboratories in your system provide DNA analyses, will you be answering this questionnaire for:

a) Your laboratory only; or for
b) All laboratories performing DNA analysis in your system

3. Organizationally, where is your laboratory located?

a) Federal
b) State
c) County
d) City
e) Private
f) Other (please describe)_____

4. Is your laboratory ASCLD/LAB accredited?

a) Yes b) No

5. What type of testing is done in your laboratory? **Please circle all types of testing performed.**

a) RFLP
b) Dot-Blots and/or D1S80
c) STRS
d) Other (please describe) _____

6. For the calendar year 1997, please estimate the number of DNA cases your laboratory analyzed and reported out.

SAMPLE / EVIDENCE RETENTION

7. Does your laboratory, as a practice, always attempt to preserve a portion of the biological specimen or extracted DNA specimen for subsequent confirmation or re-testing? **Please circle the appropriate response(s).**

a) Biological Specimens
b) Extracted DNA
c) Do not preserve as a matter of practice

8. Please estimate the percentage of DNA cases analyzed during calendar year 1997 in which key biological samples were so small that they were consumed in their entirety in DNA testing.

9.After analysis, what types of case evidence do you retain and how long do you retain it in your own laboratory?

Do you retain the original evidence? Yes No
(*Here we mean the original item or garment containing the biological stain.)

If Yes, how long? _____

Do you retain a cutting of the evidence containing the stain of interest?

Yes No

If Yes, how long? _____

10. Where is the original evidence sent after it leaves the laboratory?
(e.g., it is returned to the submitting agency)

11. How long is it likely to be kept there (if you know)?

We are aware of QA/QC standards and procedures employed by ASCLD, TWGDAM, and the DAB regarding technical reviews and audits. The following questions attempt to describe the frequency, nature, and outcomes of those reviews.

INTERNAL REVIEW / REANALYSIS

**We are interested in the percent of cases in which there is review of laboratory work by personnel within the same laboratory.

12. Please estimate the percentage of DNA cases your laboratory analyzed for calendar year 1997 in which there was an additional internal review of laboratory results (for QA/QC purposes) by your laboratory peers, supervisors, quality assurance personnel, or other members of your parent agency.

13. What is the nature of this review? (**Circle all that apply** and estimate the percentage of cases analyzed by your laboratory during calendar year 1997 receiving this type of review)

	% of Cases Receiving This Type of Review
a) Review of notes and data	_____
b) Audit of procedures	_____
c) Review of final report	_____
d) Re-testing of samples	_____
e) Other (please describe)	_____

14. If you answered (d) Re-testing of samples, affirmatively, please provide additional information about this re-testing, including how those cases are selected for re-testing and who does the retesting. (If you do not retest, skip to question #17)

15. Approximately what percentage of these re-tested results agree / disagree with the original results.

Agree _____
Disagree _____

16. If results disagreed, please describe the nature of the discrepancy.

EXTERNAL AUDITING / RE-TESTING

** We are also interested in the percent of cases examined in which there is scrutiny of your labwork by personnel outside the laboratory.

17. Please estimate the percentage of actual DNA cases analyzed during calendar year 1997 (not including proficiency tests) in which there is a review of laboratory results by an external person or organization.

18. Who has provided this review for your laboratory?

19. What is the nature of this review? **(Circle all that apply** and estimate the percentage of cases analyzed during calendar year 1997 receiving this type of review.)

This Type of Review	% of Cases Receiving
a) Review of notes and data	_____
b) Audit of procedures	_____
c) Review of final report	_____
d) Re-testing of samples	_____
e) Other (please describe)	_____

20. If you answered (d) Re-testing of samples, affirmatively, please answer the following question. Compared to the original findings, approximately what percentage did these re-tested results agree / disagree with the original results (If no re-testing occurred, skip to question #22):

Agree _____
Disagree _____

21. If results disagreed, please describe the nature of the discrepancy.

DEFENSE SCRUTINY / DISCOVERY

22. Please estimate the percentage of DNA cases you analyzed in your laboratory for calendar year 1997 in which laboratory data (e.g., laboratory notes, methods book, proficiency testing results, raw data, primary laboratory work products) beyond the laboratory report is disclosed to the defendant as a result of a discovery motion?

23. Please estimate the percentage of DNA cases analyzed for calendar year 1997 in which an analyst (from your laboratory) testified at a hearing or at a trial.

24. In what percentage of your reported DNA cases that you know of is the laboratory report and/or other laboratory data reviewed by a defense expert (e.g., a person with DNA technical expertise)?

25. Please estimate the percentage of cases analyzed in 1997 in which the evidence was subsequently re-tested by a defense expert.

26. Approximately what percentage of cases in which evidence was re-tested did these results agree / disagree with original findings:

Agree _____
Disagree _____
Don't know* _____ *were not informed of results

27. If results disagreed, please describe the nature of the discrepancy.

VOLUNTARY PARTICIPATION IN BLIND PROFICIENCY TESTING STUDY

In 1996, we surveyed your laboratory to determine if your laboratory would be willing to participate in a small-scale blind proficiency testing trial program. Our original agreements with laboratories expired in March 1998; therefore, we would like to once again ask you whether your laboratory would agree to be a potential participant in this project.

Participation means that your laboratory will be placed on a list of laboratories which might: (a) receive a fictitious case disguised to look real; (b) be asked to serve as a reference laboratory in a proficiency test; or (c) both.

1) Yes, I wish to participate. 2) No, I do not wish to participate.

If you answered Yes, please read and complete the enclosed Agreement (colored paper). The Agreement is separate and distinct from this questionnaire. We realize that administrative review may be required before a final decision can be made on whether this lab will participate or not. However, please return the survey instrument as soon as possible, and indicate to us if the Agreement Form is under review.

D-5. PHASE 2 DEFENSE ATTORNEY SURVEY INSTRUMENT

University of Illinois at Chicago
Forensic Science Program
Chicago, Illinois

DEFENSE ATTORNEY SURVEY INSTRUMENT

Basic Information

In the space below, please enter the complete name and mailing address of your organization, and the name, telephone and FAX numbers, and e-mail address of the person completing the form.

Name/address of your organization:

_____

_____

_____

_____

Name of person(s) completing survey: _____

Telephone: (___)_____
FAX: (___)_____
e-mail: _____

For the following questions, please circle the number which corresponds to your best response. If you circle OTHER please specify/explain your response in the space provided.

1. In 1997, approximately how many cases involving DNA analyses were you involved as:

a) the defense attorney of record _____
b) legal DNA consultant _____

If your answer is "none," please do not answer any additional questions. Be sure your address and telephone/fax/e-mail numbers are complete and return the questionnaire in the envelope. Thank you very much.

2. In approximately what percentage of the above cases cited in #1 did your review of the DNA evidence involve:

	% of cases
Review the DNA laboratory report only	_____
Consult with an independent expert about the laboratory report	_____
Review the underlying test results	_____
Have an independent expert review the underlying test results	_____
Have an independent lab replicate/retest the same samples tested by the prosecution	_____
Have an independent lab test additional samples (other than those tested by the prosecution)	_____
Visit the laboratory that performed the testing	_____
Other _____	_____

3. In approximately what percentage of DNA cases in 1997 that you have reviewed / had retested, were the results and/or interpretations of the original lab work viewed as problematic or questionable by your expert(s)?

4. Of cases where you have had evidence re-analyzed, in approximately what percentage have the laboratory results been notably different from the initial laboratory's results?

5. If laboratory results were different, describe the nature of the discrepancy.

6. Please rate your experience in obtaining laboratory data (e.g., laboratory reports/information, test data, QA/QC records, etc) from the prosecution on a scale of 1 to 5 with 1 being easiest and 5 being most difficult.

1 2 3 4 5

7. What has been your experience in attempting to secure funds (from the courts or other sources):

a) to have DNA evidence reviewed?
b) to have DNA evidence re-tested?
c) for travel and testimony for expert witness?

8. Do you have any other comments relevant to our objective of learning about the experience of defense counted in reviewing DNA cases?

9. If you know of other defense attorneys or experts involved in the review/reanalysis of DNA cases, could you please supply us with their names, addresses, and phone numbers and we will send them a copy of this survey.

D-6. PHASE 2 EXPERT WITNESS SURVEY INSTRUMENT

University of Illinois at Chicago
Forensic Science Program
Chicago, Illinois

DNA EXPERT WITNESS / TESTING LABORATORY SURVEY INSTRUMENT

Basic Information

In the space below, please enter the complete name and mailing address of your organization, and the name, telephone and FAX numbers, and e-mail address of the person completing the form.

Name/address of your organization:

_____

_____

_____

_____

Name of person(s) completing survey: _____

Telephone: (___)_____
FAX: (___)_____
e-mail: _____

To the following questions, please provide your best response based on your experience as a DNA expert and/or testing laboratory.

1. In calendar year 1997, in about how many cases did you review DNA test results reported by another laboratory in order to provide advice to a criminal defense lawyer?

2. In calendar year 1997, in about how many cases did you review DNA test results reported by another laboratory in order to provide advice to someone other than a defense lawyer (e.g., a prosecutor, a police agency)?

3. Please estimate the percentage of the DNA cases cited in questions #1 and #2 that you:

	% of cases
a) Perform a re-test **in your laboratory** or an item of evidence that had previously been tested at another lab	_____
b) Perform a DNA test **in your laboratory** of an item of evidence that was not previously tested	_____
c) Referred a previously tested item to **another laboratory** (not your own) for a re-test	_____
d) Referred an item that was not previously tested to **another laboratory** (not your own) for DNA testing	_____

4. In what percentage of the DNA cases cited in questions #1 and #2 did you believe that the results and/or interpretations made by the original laboratory were questionable or problematic?

5. What types of problems have you found in cases you reviewed? (Please circle your response(s) and estimate the percentage of reviewed cases in which these problems occur).

	% of cases
a) Chain of custody issues (e.g., danger of confusion, switching of samples; poor documentation of sample identity)	_____
b) Sample handling issues (e.g., danger of cross-contamination)	_____
c) Inadequate documentation of what was done	_____
d) Failure to follow proper testing protocol	_____
e) Poor laboratory practices (e.g., ignoring failure of controls or failing to use controls)	_____
f) Biased or problematic interpretation of results	_____
g) Problems related to calculations of inclusion probabilities / population genetics issues	_____
h) Other _____	_____

Please elaborate on questionable or problematic issues you have seen.

6. In the cases involving re-testing (noted question #3a and #3c above), please estimate the percentage in which you obtained different test results from the initial crime laboratory's results.

If any results were different, please describe the nature of the discrepancy.

7. In the cases involving testing of an additional item(s) (noted in question #3b and #3d above), please estimate the percentage in which the additional testing produced important or unexpected results (e.g., evidence of additional or alternative perpetrators).

Please explain the nature of any such results

8. In your experience, how often have you found that critical biological evidence was consumed in its entirety by the government laboratory in its initial DNA testing?

9. In cases where you have been asked to review DNA test results of another laboratory, have you always had access to the materials you believed you needed to perform an adequate review?

If not, please answer the following:

a) What materials did you believe you needed that were not provided?

b) What was the reason (if you know) that the materials you believed you needed were not provided?

10. If you know of other defense attorneys or experts involved in the review/reanalysis of DNA cases, could you please supply us with their names, addresses, and phone numbers and we will send them a copy of this survey.

Thank you very much for your responses. Please return this survey in the enclosed postage paid envelope.

E-1. AGREEMENT WITH LABORATORIES

Blind DNA Proficiency Testing Feasibility Project
Agreement to Participate

Name of Laboratory: _____

** If you have indicated that you would like your laboratory to be in a pool of DNA typing labs that may be utilized as trial "blind" test sites, or as reference laboratories, please complete this agreement and return it to us. We will return a fully signed copy to you. **

The above-named laboratory agrees to be a participant in a NIJ-funded blind DNA proficiency testing (PT) feasibility project being conducted by the University of Illinois at Chicago (UIC). Agreement to participate means that the above-named laboratory, during the period of March 1, 1999 to August 31, 1999: (a) may be asked to be a reference laboratory for the DNA typing of certain specimens; and/or (b) may be selected and used as a trial blind testing site.

To test the feasibility of blind DNA PT, specimens will be introduced into trial-site laboratories in the guise of real cases, complete with all the appropriate and expected evidence containers, numbers, forms, paperwork, and so forth. This introduction will be effected with the cooperation of a law enforcement agency, and possibly the prosecutor's office.

With respect to the actual trial blind PT feasibility testing, UIC project directors give you assurances that:

1) specimens in the fictitious cases will be manufactured following the "Guidelines for DNA Proficiency Test Manufacturing and Reporting" (Crime Lab. Digest 21(2,Apr):27-32, 1994); specifically, they will be manufactured by a TWGDAM-approved PT manufacturer, and biological specimens employed will come from donors who have tested Negative for HIV, Hepatitis B and Hepatitis C by ELISA..

2) once your DNA testing and reporting have been completed, and the submitting agency notifies our project office, we will notify you. We will tell you that the case in question was a blind PT, and we will tell you how your results compared with those of our reference laboratories.

3) we will not reveal which specific laboratories or specific examiners participated in any trial testing in this project unless we are legally required to do so; we will not indicate the identities of participating laboratories in our project write-ups.

4) your participation in this project is totally voluntary, and you may discontinue your involvement at any time.

By agreeing to participate in the project, you agree to

1) contact the project office if you suspect that a submitted case is really a "blind" PT; we will inform you if your suspicion is correct.

2) keep confidential the fact of your involvement as a trial test site and/or as a reference laboratory, unless you are legally required to reveal your participation.

3) purge from your files (including computer-stored records), if possible and permissible, any and all records connected to a "case" introduced into your lab as a trial blind PT or clearly identify the records retained as being connected to a blind PT; and to assist us to the extent possible in causing such records to be purged from police and prosecutor files to the extent possible and permissible, or to be clearly identified as being connected to a blind PT, once we have revealed to you that the "case" was in fact a trial blind PT.

4) completely purge from your databases, and cause to have purged from any centrally-maintained databases, any and all DNA types and profiles that were entered into such databases as the result of analysis in a "case" that we reveal to you was fictitious.

5) cooperate with the UIC project team in analyzing the results and problems, and assessing the costs, of conducting a blind PT, if a blind PT was introduced into your lab and went through undetected, or if a blind PT was introduced into your lab and was detected.

Responsible Official for the laboratory:	Project Director for the UIC:
_____	_____
Signature	Joseph L. Peterson, or R. E. Gaensslen
_____	
Typed Name	
_____	
Title	

DNA Section Examiners: [This item was made optional in Phase 2]

_____	_____
Signature	Printed or Typed Name
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

** Examiners signatures indicate that the laboratory's possible participation in this project has been explained to them, and that this participation may involve the deception of one or more examiners in a trial blind PT.

E-2. AGREEMENT WITH LAW ENFORCEMENT AGENCIES

Blind DNA Proficiency Testing Feasibility Project
Law Enforcement Agency Agreement to Participate

Law Enforcement Agency: _____

The above-named law enforcement agency agrees to be a participant in a NIJ-funded blind DNA proficiency testing (PT) feasibility project being conducted by the University of Illinois at Chicago (UIC). Agreement to participate means that the above-named agency, during the period of March 1, 1999 – August 31, 1999, may be asked to assist the project team in introducing manufactured biological evidence specimens, disguised as a legitimate case submission, to a forensic-science laboratory in your jurisdiction for DNA typing.

To test the feasibility of blind DNA PT, specimens will be introduced into trial-site laboratories in the guise of real cases, complete with all the appropriate and expected evidence containers, numbers, forms, and related paperwork. This introduction will require the cooperation of a law enforcement agency that normally submits physical evidence to that laboratory for analysis. For these blind PT test introductions to serve as an honest measure of blind PT feasibility, the participating law enforcement agency, and its participating designated personnel, must make convincing written and oral statements (as needed) to the laboratory and its personnel to insure that the laboratory believes the proficiency test is a real case. It is understood that these representations by law enforcement agencies and agents constitute temporary deception of the laboratory and its staff. It is also to be noted that any laboratory to be tested in this project has agreed in advance to the procedures described here. Successful blind PT test introduction may also require the cooperation of other entities in the jurisdiction, such as a prosecutor's office, or a hospital emergency-room staff.

With respect to the actual trial blind PT feasibility testing, UIC project directors give you assurances that:

1) The laboratory that receives the fictitious case has agreed in writing to be a participant in this project, and understood in so agreeing that a fictitious case containing manufactured evidence could be submitted to that lab for DNA typing.
2) Specimens for the fictitious cases will be manufactured following the "Guidelines for DNA Proficiency Test Manufacturing and Reporting" (Crime Lab. Digest 21(2,Apr):27-32, 1994); specifically, they will be manufactured by a TWGDAM (Technical Working Group on DNA Analysis Methods)-approved PT manufacturer, and all biological specimens employed will come from donors who have tested Negative for HIV, Hepatitis B and Hepatitis C by ELISA.

3) Once the DNA testing and reporting of results have been completed, the laboratory notifies the submitting agency by issuing a written report, and the agency notifies the UIC project directors, we will notify the laboratory that the case was a blind PT. We will also tell the laboratory how its results compared with those of our reference laboratories.
4) We have agreed with participating laboratories to disclose to them that a "case" is a blind PT if the lab suspects that it is, and if in fact it is, and if they contact us and ask whether it is. If this contingency occurs, we will notify you as well.
5) We will not reveal which specific laboratories or specific examiners participated in any trial testing in this project unless we are legally required to do so; we will not indicate the identities of participating laboratories in our project write-ups. Likewise, we will not reveal which specific law enforcement agencies cooperated with us in any blind trial PT unless we are legally required to do so; and we will not indicate the identities of participating law enforcement agencies in our project write-ups.
6) Your participation in this project is totally voluntary, and you may discontinue your involvement at any time.

By agreeing to participate in the project, you agree

1) Not to disclose to anyone connected with the forensic-science laboratory your cooperation with the project team in helping to submit a blind PT DNA "case".
2) To assist UIC project staff in the creation of the proficiency test and related case materials and to make as convincing a portrayal as possible to the laboratory that the case is genuine
3) To keep confidential the fact of your involvement as a participant in the project, unless you are legally required to reveal your participation.
4) To notify UIC Project Directors when the agency receives the laboratory report
5) To either (i) purge from your files (including computer-stored records), if possible and permissible, any and all records connected to a "case" introduced through your agency as a trial blind PT; or (ii) identify, mark or flag the records retained as being connected to a blind PT of the laboratory
6) To cooperate with the UIC project team in analyzing the results and problems, and assessing the costs, of conducting a blind PT in the manner described, regardless of whether a blind PT was introduced into a forensic-science lab through your agency and went through undetected, or was introduced and detected.

Responsible Official for the law enforcement agency:	Project Director for the UIC:
_____	_____
Signature	Joseph L. Peterson or R.E. Gaensslen
_____	
Typed Name	
_____	
Title	

Law Enforcement Agency Personnel Actually Participating:**

Signatures	Printed or Typed Names

** Actual participating personnel signatures indicate that the agency's possible participation in this project has been fully explained to them, and that this participation will involve the deception of one or more individuals in the forensic-science laboratory during a trial blind PT. These signatures are not required by us, and may be left blank at your discretion.

E-3. AGREEMENT WITH CONDUIT LABORATORIES

Blind DNA Proficiency Testing Feasibility Project
Agreement to Participate as a Conduit Laboratory

Name of Laboratory: _____

The above-named laboratory agrees to be a participant in a NIJ-funded blind DNA proficiency testing (PT) feasibility project being conducted by the University of Illinois at Chicago (UIC). Agreement to participate means that the above-named laboratory, during the period November 1, 1998 – August 1, 1999, may be asked to act as a submitting "conduit" lab in transmitting certain specimens to another laboratory that has agreed to serve as a trial blind feasibility testing site.

To test the feasibility of blind DNA PT, specimens will be introduced into trial-site laboratories in the guise of real cases, complete with all the appropriate and expected evidence containers, numbers, forms, and paperwork. This introduction will be effected in some instances with the cooperation of a law enforcement agency, and possibly the prosecutor's office, and in others with the cooperation of a second "conduit" laboratory.

With respect to the actual trial blind PT feasibility testing, UIC project directors give you assurances that:

1) Specimens in the fictitious cases will be manufactured following the "Guidelines for DNA Proficiency Test Manufacturing and Reporting" (Crime Lab. Digest 21(2,Apr):27-32, 1994); specifically, they will be manufactured by a TWGDAM-approved PT manufacturer, and biological specimens employed will come from donors who have tested Negative for HIV, Hepatitis B and Hepatitis C by ELISA.

2) The target-site laboratory has agreed explicitly, in writing , with us to be a potential blind PT target site.

3) Once DNA testing and reporting have been completed, and you so notify our project office, we will tell the target-site laboratory that the case in question was a blind PT, and we will also tell them how their results compared with those of our reference laboratories.

4) We have agreed to inform the target-site laboratory that a submitted "case" is really a blind PT if they have such suspicions and ask us explicitly. Under such circumstances, we will also inform you of these facts.

5) we will not reveal which specific laboratories or specific examiners participated in any trial testing in this project unless we are legally required to do so; we will not indicate the identities of participating laboratories in our project write-ups.

6) your participation in this project is totally voluntary, and you may discontinue your involvement at any time.

By agreeing to participate in the project, you agree to

1) Contact the project office if the target-site laboratory suspects that a submitted case is really a "blind" PT, and poses this question to you.

2) Keep confidential the fact of your involvement as a "conduit" laboratory, unless you are legally required to reveal your participation.

3) Either (i) purge from your files (including computer-stored records), if possible and permissible, any and all records connected to a "case" introduced by your laboratory into another lab as a trial blind PT; or (ii) clearly identify, mark or flag the records retained as being connected to a blind PT; and to assist us to the extent possible in causing such records to be purged from police and prosecutor files to the extent possible and permissible, or to be clearly identified as being connected to a blind PT, once we have revealed to the target-site laboratory that the "case" was in fact a trial blind PT.

4) Completely purge from your databases, and cause to have purged from any centrally-maintained databases, any and all DNA types and profiles that were entered into such databases as the result of analysis in a "case" that we have manufactured, and that we have revealed to the target-site laboratory as fictitious.

5) Cooperate with the UIC project team in analyzing the results and problems, and assessing the costs, of conducting a blind PT, if a blind PT was introduced through your lab, regardless of whether the case went through undetected, or was introduced through your lab and was detected.

Responsible Official(s)
for the "conduit" laboratory:

UIC Project Directors:

_____
Signature

_____
Joseph L. Peterson or R.E. Gaensslen

_____
Typed Name

_____
Title

F-1. INSTITUTIONAL REVIEW BOARD (IRB) MATERIALS

I. Introduction about Protection of Human Research Subjects

The Institutional Review Board (IRB) reviews research to ensure that the federal regulations for protecting human research subjects outlined in both the Department of Health and Human Services (HHS) regulations (45 CFR 46) and the Food and Drug Administration (FDA) regulations (21CFR Parts 50 & 56) as well as other requirements are met. The University of Illinois at Chicago (UIC) Multiple Project Assurance (MPA) (M 1095) awarded by the Office for Project from Research Risks (OPRR) at the National Institute of Health, is a written pledge to follow federal guidelines for protecting human research subjects. Under the MPA, the University officials, investigators, the IRB, and the OPRS staff work together to comply with those guidelines. Investigators have responsibilities when conducting research involving human subjects.

II. UIC Institutional Review Board Application

1. Research Introduction

Funded in 1996, this study sought to test the feasibility of external, blind DNA proficiency testing in forensic science laboratories in the U.S.. Proficiency testing, the submission of samples to laboratory scientists to determine the accuracy of their testing procedures, is the cornerstone of quality assurance programs in many fields. Although most forensic DNA laboratories employ "open" proficiency testing as a means of quality control, "blind" proficiency tests - where the analysts are unaware the sample they are examining is a "test", are advocated by some to be a better test of an examiner's ability. In the DNA Identification Act of 1994, the U.S. Congress directed the National Institute of Justice to sponsor research to evaluate the feasibility of a "blind" DNA proficiency testing program. In a competitive process, UIC's proposal was selected to investigate this question. To accomplish its goal, this project established a National Forensic DNA Review Panel of acknowledged experts to guide its work, reviewed the relevant clinical and forensic testing literature, surveyed law enforcement agencies and forensic DNA laboratories regarding their evidence collection, examination, and quality assurance practices, and surveyed independent experts and defense attorneys concerning their re-review and legal examination practices. But most importantly, it proposed to conduct limited blind proficiency tests on selected voluntarily participating forensic DNA laboratories.

Human subjects were involved in three ways: First, our proficiency test subcontractor, the American Registry of Pathology, recruited biological sample donors from its staff to produce proficiency test samples. Donors were informed of the potential benefits and risks. Second, during the blind proficiency tests, laboratory analysts examined the submitted test evidence samples using their routine protocols. Only forensic laboratories and analysts who had agreed to submit to blind proficiency testing were used. Third, forensic DNA laboratories, defense attorneys and expert witnesses were surveyed to assess current quality control, auditing, and legal review procedures used to evaluate such evidence.

All laboratories and analysts agreeing to participate in this study were advised that their identities would not be revealed nor associated with study results, that they were free to discontinue their involvement in the study at any time, and that they would experience no other adverse consequences. With respect to the donors of biological specimens, we explained the use of their specimens in the creation of fictitious cases, and that although their names would not be revealed, their DNA types were, nonetheless, highly individualistic and might be entered into law enforcement DNA databases. These potential donors (in Phase II) were also informed that the Code of Federal Regulations authorized us to protect their identities and that all DNA types entered into databases as a result of this research would be expunged. (Note: Because we had not obtained this opinion when samples were obtained in Phase I, the initial consent form did not contain this proviso.)

Because blind proficiency testing was being considered by the U.S. Congress as a requirement for forensic laboratories to receive federal assistance, it was imperative to supply Congress with the results of feasibility testing. Furthermore, this research benefits the forensic community by assessing the feasibility of this mechanism for evaluating the reliability of forensic DNA testing. All known forensic DNA laboratories in the U.S. were contacted for this study. Only laboratories and analysts that agreed to participate in the research were considered for blind proficiency testing. As for surveying defense attorneys and expert witnesses, due to the specialized and limited sample population, a snowball sampling technique was used to generate names of potential subjects.

2. The Tasks/Tests or Procedures Subjects Were Asked to Complete

A. Preparation of Blind Proficiency Test Specimens:

A total of eight biological specimen donors were recruited by our subcontractor from among staff and other healthy adults known to them to have been willing to be biological specimen donors for other proficiency testing projects. IRB approved informed consent forms were provided both to the subcontractor and to the volunteer donors (Appendix F-2).

B. Subjects of Blind Proficiency Tests:

1) Forensic DNA Laboratories: Blind proficiency tests were submitted to 15 participating laboratories (and 7 reference laboratories) through routine channels as if they were regular evidence. Upon completion of the test, laboratories were notified the samples were proficiency tests, and were then asked to purge all results from relevant databases (Appendix I).

2) Laboratory Analysts: Individual analysts of selected participating laboratories performed routine forensic DNA analysis on submitted blind proficiency tests.

C. Subjects of Surveys:

1) Forensic DNA Laboratories: Two IRB approved surveys were distributed to forensic laboratories, during Phases I and II of the study. In both instances, laboratories were asked to complete the survey according to instructions and return the survey to the UIC researchers. (Appendices D-1, and D-4).

2) Defense Attorneys: Subjects were asked to complete the IRB approved mail survey and return it to UIC researchers (Appendix D-5).

3) Expert Witnesses: Subjects were asked to complete the IRB approved mail survey and return it to UIC researchers. Based on the content of self-reports, selected experts were approached for in-depth telephone interviews (Appendix D-6).

3. Recruitment Procedures

A. Blind Proficiency Test Sample Donors:

Subjects were either staff members of the American Registry of Pathology who volunteered to donate samples for the manufacturing of blind proficiency tests or other healthy adults known to the American Registry of Pathology staff to have been donors for other proficiency test projects. The American Registry of Pathology routinely engages in the preparation of declared proficiency tests under its Forensic Identity Program that is available to clinical and forensic laboratories, and consequently employs standard/clinically approved procedures for gathering such samples.

B. Selection of Subjects for Blind Proficiency Tests:

1) Forensic DNA Laboratories: Our goal was to contact every laboratory in the U.S. performing forensic DNA analysis. A list was compiled based on the National Institute of Justice DNA grantee list, the FBI's 1995 Combined DNA Index System (CODIS) Survey laboratory register, and telephone inquiries.
2) Laboratory Analysts: Laboratories and analysts that expressed interest in the blind testing program were issued agreement forms that outlined the procedures to be followed. Laboratories and analysts executing such agreements were placed on a list of laboratories eligible to receive the blind tests.

C. Selection of Subjects for Surveys:

1) Forensic DNA Laboratories: Same as above.
2) Defense Attorneys and Expert Witnesses: Due to the highly specialized and limited population of witnesses and attorneys with DNA expertise, the National Forensic DNA Review Panel provided names for the initial sample list, and then a snowball sampling technique was used to expand the sample pool.

4. Criteria for Inclusion and Exclusion of Subjects in the Study

A. Biological Specimen Donors:
Staff of the American Registry of Pathology and other healthy adults who have previously provided specimens for other proficiency tests were included. Those not wishing to participate were excluded.

B. Forensic Laboratories:
If the laboratories indicated a willingness to participate in the blind testing portion of the project, they were asked to supply the researchers with a list of law enforcement agencies that routinely submitted biological evidence to them for DNA testing. They also identified other nonDNA typing forensic laboratories in their jurisdiction that would also routinely submit samples for DNA testing. We next contacted those law enforcement agencies and laboratories to determine which would agree to work with us in preparing the blind "evidence" test sample to the laboratory (Appendices D-2 and D-3). The actual blind proficiency tests were, therefore, issued to that smaller subset of laboratories that indicated a willingness to participate and for which we were able to identify a law enforcement agency or conduit laboratory willing to cooperate with us in submitting the sample. The laboratory's decision to participate meant that a "test" case may be submitted to it within a defined time period. Laboratories were assured their involvement was voluntary, could be discontinued at any time, and that laboratories and individual analysts would not be identified in subsequent reports or publications. Furthermore, if at any point in the testing procedure a laboratory analyst or supervisor became suspicious that the case before them was a test, they were to contact the project staff. Laboratories that indicated in our initial letter/survey inquiry that they did not wish to participate in the study, as well as those that did not supply the necessary signatures on the agreement forms, were excluded from the blind proficiency testing. In addition, laboratories for which we were unable to locate an agency or laboratory that would submit the sample were excluded.

C. Subjects of Surveys:
1) This study surveyed all known forensic laboratories in the U.S. performing DNA typing and 68% responded with survey information. Only subjects who replied to surveys were included. Those not replying to the surveys were excluded.
2) Defense Attorneys and Expert Witnesses: Due to the highly specialized and limited sampling population, members of the National Forensic DNA Review Panel provided names for the initial sample list, and then a snowball sampling technique was used to expand the sample pool. Eleven expert witnesses and six defense attorneys returned survey results.

5. Risks, Deception, and Benefits of the Research

A. The risks to the subjects (specimen donors, forensic laboratories, analysts, defense attorneys, and expert witnesses) were minimal.
1) Biological Specimen Donors:
We minimized the risk to specimen donors by 1) Using a laboratory (the American Registry of Pathology) and personnel experienced in the drawing of biological samples (using a trained phlebotomist) and the creation of proficiency tests;

2) Constructing a detailed consent form (Appendix F-2) which fully explained the process by which the samples would be analyzed and (potentially) data based by the DNA typing laboratories in the research; 3) Warning the donor not to supply a sample if they had any concerns over having their DNA profile included temporarily within a federal, state or local DNA database; 4) Obtaining assurances from participating laboratories that they would extract all databased DNA types at the completion of testing; and 5) Obtaining an assurance from the U.S. Department of Justice that 28 C.F.R. pt. 22 would enable UIC and the subcontractor to resist any judicial administrative or legislative proceeding (such as a subpoena) process for obtaining the identity of a person submitting a biological sample without their written consent.

2) Testing Laboratories:

In terms of the blind proficiency testing, agreements to participate were secured, test samples were presented to the laboratories as "routine evidence submissions", and were processed through normal procedures using their standard procedures. Therefore analysts were expected to perform standard DNA typing procedures for which they had been trained and which they performed on case evidence on a routine basis. Laboratories and analysts were assured they would not be identified. All blind proficiency specimens were tested for the presence of viral markers (hepatitis, HIV) in accordance with FDA standards to protect the forensic analysts who would be subsequently analyzing the sample. In addition, the Guidelines for DNA Proficiency Test Manufacturing and Reporting approved by the forensic community and which set standards for commercial proficiency test providers were followed.

3) Survey Respondents:

Respondents were free not to respond to the various survey instruments issued during the course of the study and no laboratory organizations, law enforcement agencies, or individuals would be identified or associated with their responses in publication of final results and recommendations.

B. The purpose of the study was to ascertain the feasibility of conducting external, blind DNA proficiency testing and to assess the associated costs and benefits. Therefore, "deception" was essential in this study. Blind DNA Proficiency test samples were introduced into voluntarily participating forensic laboratories. The test samples were disguised as actual cases. However, laboratories must have elected to be participants in order to receive such samples, and were informed they would be receiving such a disguised case at some point within a defined time period. Because laboratories were informed of study procedures, and had agreed to participate before the samples were issued to them, the risks associated with this "deception" were minimal.

C. Individual forensic examiners and the profession-at-large shall benefit from this study to determine the feasibility of such testing. Many professionals and policy makers have called for the introduction of blind DNA testing as a method for insuring high quality control (QC) standards in the field.

6. Confidentiality of Data

A. Provisions Made to Maintain Confidentiality of Data:

Only the Principal Investigators and Research Assistants had access to the raw data. Raw data was not made available to anyone other than the principal investigators and research assistants. Data provided to the study advisory panel and, ultimately, in the final reports contained no identifiers.

B. The Location Where the Data Is Kept

1) Biological Donor Samples:
   The American Registry of Pathology agreed to destroy biological samples after preparation of the tests. All records were deleted from databases upon the completion of the blind tests.
2) Blind Proficiency Test Data:
   Laboratory proficiency test results are kept in the locked office of Principal Investigator.
3) Survey Data:

Survey results are kept in the locked office of the Principal Investigator. Analyzed data are stored on password protected computer inside locked office of the Principle Investigator. Surveys will be destroyed 3 years after the completion of the project, in accordance to NIJ standards.

III. The Dates the Project Was Approved By UIC Institutional Review Board

1. April 2, 1996. IRB No. H-96-062.
2. April 2, 1997. IRB No. H-96-062.
3. April 2, 1998. IRB No. H-96-062.
4. April 19, 1999. IRB No. H-99-194.
5. January 10, 2000. IRB No. H-99-194.

F-2. Biological Specimen Donor Informed Consent Form

University of Illinois at Chicago
Blind DNA Proficiency Testing Project

BIOLOGICAL SPECIMEN DONOR CONSENT FORM

I hereby give my consent to donate a biological specimen or specimens to be used in the National Institute of Justice funded study of the feasibility of blind DNA proficiency testing for the nation's forensic DNA testing laboratories. The objective of this study is to prepare realistic biological evidence items, such as would be encountered in routine crime laboratory casework, to see if such "evidence" can be submitted to forensic DNA testing laboratories through normal channels without the laboratory recognizing these items are actually a proficiency test. If such procedures are found feasible, they may prove to be important in programs to assure that such laboratories perform the highest quality analyses.

The University of Illinois at Chicago (UIC) is the primary grantee for this project. The UIC Project has contracted with the American Registry of Pathology (ARP) to manufacture biological specimens for use in blind proficiency testing. Participation in this study is voluntary and your refusal to participate will involve no penalty or loss of any benefits to which you are otherwise entitled. Once you donate biological specimens you will have 14 days to notify the AFIP representative that you have decided you do not want your specimens included in this study. After 14 days, you will not be able to revoke your decision to participate.

You may choose to be a donor of blood, semen, vaginal swabs and/or oral (buccal) swabs. The biological specimens will be used to make stains on clothing or other materials to simulate the "evidence" gathered in a fictitious violent or sexual assault crime. The blood may also be used to simulate blood taken from a fictitious "suspect" or "victim". After preparing the stains, the AFIP laboratory will retain any remaining portions of your sample until the conclusion of the project. At that time, the samples will be destroyed.

Any names or identifiers used in these simulated crime cases will be fictional. Your name will not be used or revealed to anyone. You should be aware, however, that even though your name will not be used, the DNA profile derived from your specimen is highly individualistic. The DNA testing to be performed on your specimen in this study is for identification purposes only and will not be used for any other clinical, medical or genetic purposes. The Project Directors are nevertheless mindful of the concern the public has about any type of DNA testing and will take every precaution to protect this sensitive information.

If you donate blood, a blood specimen will be taken by means of phlebotomy and you will take full responsibility for the minimal health risks associated with this procedure. You also understand that this specimen will be tested for hepatitis B, hepatitis C and human immunodeficiency virus (AIDS) markers. Any positive results will be brought to your attention.

The DNA testing laboratories that receive these simulated cases will determine the DNA types of the biological "evidence," and compare types from the "evidence" with those of "victims" and "suspects" in the fictitious case.

Besides making the above comparisons, it is a routine for some forensic DNA laboratories to enter the DNA types of biological evidence into local, state, and possibly a national computer data base. These databases are used by forensic laboratories to determine if biological specimen DNA types in a criminal investigation "match" the DNA types from any other cases, and if the DNA types match those of known offenders. Many states have passed laws that require persons convicted of certain serious offenses submit blood samples to authorities for DNA typing and entry into data bases.
Even though none of the participating laboratories will know the true identity of the specimen donor, because of the possibility of your DNA profile being entered into a DNA database we want to fully disclose to you the following possibilities. We consider these possibilities to be exceedingly remote, and thus the potential risk to be very small.

Because the forensic testing laboratory will presumably believe this to be an actual case, your DNA profile may be temporarily entered into one of these DNA databases to determine if it corresponds to any of the DNA profiles already in the data base, either from known convicted offenders, or from other unsolved crimes in that locality, state or throughout the country. We do not know with certainty how long your DNA profile might be in one or more of these DNA databases, but it could range from a matter of days up to several months.

As soon as the laboratory reports its results to the law enforcement agency which submitted the fictitious case "evidence," the laboratory will be informed immediately that the case they just worked was a fictional case/test and not real evidence. The project directors have obtained written assurances from the laboratories participating in this research that they will immediately remove any and all DNA types and profiles from this fictional case that may have been entered into any DNA computerized data bases. Laboratories will also be requested to destroy the actual proficiency test stains and any subsamples or DNA prepared from them at the conclusion of the project.

The Project Directors will also request the laboratory to send us a written statement verifying your DNA profile has been removed from the relevant computerized data bases, and that ALL proficiency-test stain specimens have been destroyed. You should be aware, however, that although we have agreements in advance and that we further seek assurances that any donor's DNA profile has been deleted, there is a remote chance this deletion will not occur and your DNA profile could be left in the file.

In addition, if you have reason to believe that your DNA profile from some past, present or future act may be included in a law enforcement DNA data base, you should not participate in this study. We issue this warning for the following reasons: if you have previously been convicted of a crime in a state that has a DNA databanking law, and that law required you to donate a blood specimen for DNA databanking, a "match" may occur between that existing profile and the DNA profile derived from this new fictitious case. Secondly, if, while your DNA profile from this fictional case is being stored in a data base, you become involved in some act that would lead a law enforcement agency to enter your DNA profile into a data base, your DNA profile from the proficiency test specimen might be matched with the DNA profile from the real case and the police may seek your true identity. If such a "match" does occur between the DNA derived from a criminal inquiry and the DNA in this fictional case, law enforcement authorities could seek to determine your true identity from the ARP which obtained your biological specimen. We are authorized, however, by the Code of Federal Regulations to protect your identity. UIC, as a primary grantee, and ARP, as biological specimen subcontractor, will not reveal your identity, nor may such information be used in any judicial proceeding, without your written consent.

My signature on this form, giving consent to be a donor of a specimen (or specimens) for preparation of blind proficiency-testing samples, indicates that: i) I have read and understood the explanations of the benefits and risks of this donation; ii) my consent is given freely and without coercion; and iii) any questions I asked were answered.

Should you have any questions about the research and rights of research subjects, you may contact either James Canik, representing the ARP contract laboratory (301) 319-0210, or either Prof. Joseph L. Peterson or Prof. Robert E. Gaensslen, Co-Principal Investigators, University of Illinois at Chicago, (312) 413-0439. If you have any additional concerns about your rights as a participant in this study, you can contact the UIC Office for Protection from Research Risks at (312) 996-9299.

_____ _____

Signature Date

_____

Printed or Typed Name

Revised 3/31/99

G. EXAMPLE SPECIFICATIONS FOR MANUFACTURING OF BIOLOGICAL EVIDENCE/CASES

There are two REFERENCE Labs: (names / addresses given)

There are four DONORS total, two males and two females. They are referred to as Male 1, Male 2, Female 1, and Female 2.
Cases x through x will be manufactured using Male 1 and Female 1.
Additional cases will be manufactured later using Male 2 and Female 2.
For now, only specimens from Male 1 and Female 1 will be collected.

CLOTHING from Donors
If possible, obtain a set (under and outer) of well-worn clothing from each donor. Underwear from Female donors is especially important. It should be thoroughly worn, and not washed before donation to you; it should not be worn within 96 hours after sexual intercourse, nor during menstrual period.

BIOLOGICAL EVIDENCE Collection from Donors:

Self-collection of vaginal swabs by Female donors (general): Do not collect during menstrual period (i.e. should be blood-free). Do not collect within 96 hrs of intercourse. Collect four swabs at a time. Wait a day or so (min 12 hrs) in between collections.

Self-collection of semen specimens by Male donors (general): Collect specimen into a clean (sterilized if possible) container, and store in a frig until delivered to you. Specimen should not be contaminated with any other body fluids, or other materials (such as lubricants). Specimen must not be contaminated with body fluids from any other person.

REFERENCE and ARCHIVE

Collect three purple-top vacutainer (EDTA) tubes, approx 7 mL, from Male 1

Collect three purple-top vacutainer (EDTA) tubes, approx 7 mL, from Female 1

ONE tube from each donor used to aliquot archival known bloodstains, retained by you.

ONE tube from each donor, labeled "MALE 1", "FEMALE 1" sent asap by FedEx in styrofoam container with cold-paks or ice to CSP reference lab.

ONE tube from each donor, labeled "MALE 1", "FEMALE 1" sent asap by FedEx in styrofoam container with cold-paks or ice to OSP reference lab.

ONE vaginal swab self-collected from Female 1, to which has been added at least 50 µL self-collected semen from Male 1, and then thoroughly dried (in the same manner as described below for individual cases) sent to xxx Reference Lab.

ONE vaginal swab self-collected from Female 1, to which has been added at least 50 µL self-collected semen from Male 1, and then thoroughly dried (in the same manner as described below for individual cases) sent to xxx Reference Lab.

ONE vaginal swabs self-collected from Female 1, air dried, retained by you (Archive)

200 µL semen from Male 1, frozen at either -20° or -70° , retained by you (Archive)

Summary of Specimens needed from the donors (required for manufacture of case evidence)

Female 1: three purple-top vacutainer (EDTA) tubes (Ref / Archive)
one red-top tube of blood (approx 7 mL) (Case x)
purple-top (EDTA) tube sufficient to yield 1-2 mL (Case x)
purple-top (EDTA) tube sufficient to yield 1 mL for blot card AND 15-20 mL for spatter/smear onto shirt (Case x)
purple-top (EDTA) tube sufficient to yield 2-3 mL for bloodstain card (Case x)
three swabs using any available (Ref / Archive)
pair of swabs using xx kit swabs (Case x)
pair of swabs using xx-consistent kit swabs (Case x)
set of well worn clothing; esp. panties, underclothing
(One pair of panties for Case x)
pair of swabs using xx-consistent kit swabs (Case x)

Male 1: three purple-top vacutainer (EDTA) tubes (Ref / Archive)
one red-top (clot) tube (Case x)
one purple-top tube (Case x)
purple-top (EDTA) tube sufficient to yield 1-2 mL (Case x)
purple-top (EDTA) tube sufficient to yield 1-2 mL (Case x)
purple-top (EDTA) tube sufficient to yield 1 mL (Case x)
purple-top (EDTA) tube sufficient ot yield 2-3 mL (Case 4x
semen specimen 300-350 µL for swabs (Ref / Archive)
semen specimen 100-150 µL (Case x)
semen specimen 100-150 µL for swabs + 100-200 µL for panties (Case x)
semen specimen 100-150 µL (Case x)

CASE x. Sexual Assault Case - will be submitted for DNA typing to xxxxxxxx by xxx xxxxxxxxxxxxxx Police, Forensic Science Laboratory, Attn: xxxxxxxxxxxxxxxxxxxxxx

Supplied to you:

1. One pair (2) vaginal swabs in factory-sealed packaging (from xx standard kit); assemblable cardboard box for swabs after collection (from xx kit)
2. One red-top vacutainer tube, with bubble-wrap container (from xx kit)
3. Cotton cloth swatch, about 2.5 x 4 inches
4. Cotton cloth swatch, about 2.5 x 4 inches

Case assembly:

1. Female 1 self-collect semen-free, blood-free vaginal swabs using both vaginal swabs described in #1 above - follow general guidelines for swab collection. Allow swabs to air dry. Carefully return swabs to factory-supplied paper container. Provide swabs to you.

2a. Collect blood by venipuncture from Female 1 into red-top vacutainer described in #2 above, sufficient to fill tube.

2b. Mix up the blood in the vacutainer; then pipette out blood onto the 2.5 x 4 inch cloth swatch to form a "circle" of bloodstain. Cover a significant portion of the cloth, but leave white cloth space at the ends for writing. Allow to air-dry. This will be "victim" blood/bloodstain. The tube can be labeled with name of victim (Lname, FName), date, time and initials of phlebotomist (do not have to be the real initials). The date on this tube should pre-date the date on the "suspect" tube by at least 10 days.

3a. Collect blood by venipuncture from Male 1 into purple-top vacutainer, sufficient to yield 2-3 mL.

3b. Mix the blood in the vacutainer; then pipette out blood onto the 2.5 x 4 inch cloth swatch to form a "circle" of bloodstain. Cover a significant portion of the cloth, but leave white cloth space at the ends for writing. Allow to air-dry. This will be "suspect" bloodstain. Its container can be labeled as "suspect" / "Lname, FName"

4. Have Male 1 self-collect semen specimen into a clean (sterilized if possible) container, and store in a frig until delivered to you - follow general guidelines stated above.

5. Mix up Male 1 seminal fluid specimen to insure maximal homogeneity; then add a quantity of semen sufficient to contain at least 200,000 sperm cells (300 - 500 ng DNA) [figuring about 2.5 pg DNA / cell] to each swab collected from Female 1. Allow swabs to air dry. Package in white assemblable cardboard box provided with kit.

FOUR items (pair of swabs is considered one item) will be supplied to the submitting agency for this case.

Case 1 "Scenario"

State of xx v. LName, FName

Fname LName (WM, DOB 4/5/69) is accused of sexually assaulting Fname LName (WF, DOB 8/23/65) in her home in ------, -- on July 17, 1997. In this case, victim knew who the alleged perpetrator was, though she had not had any personal contact with him prior to the incident. Suspect denies the assault, and any contact with the complainant.

Victim called police shortly after incident. Taken to hospital ER. Kit taken.

Suspect located several days later, questioned, and arrested. Exemplar specimens taken pursuant to court order several days after arrest.

Circumstances of this case / investigative information / suspect's background indicate that a search of the CODIS database for this suspect's profile may be indicated.

Transmittal to Agency / Conduit Lab - With Biological Evidence Specimens

XXXX Laboratory, ATTN xxxxxxx

Included in this package are: (1) this page, indicating what is being shipped, and instructions for completing the submission of the "case" to the DNA lab; (2) a copy of our manufacturing specifications for this "case"; and (3) a brief "scenario" around which the "case" was constructed.

You should receive in this package (from the blind PT test manufacturer):

Pair of vaginal swabs, marked: "Fname Lname" / BT Female 1
[each swab spiked w/ at least 200,000 sperm cells (300 - 500 ng DNA)

Red-top tube from xx kit, marked: Lname, FName / July 17, 1997 / 9:30 p.m. / L.M.

Blood from this tube was used to make the "victim" bloodstain on cotton cloth.

One approx. 2.5 x 4 inch cloth swatch with a "circle" of bloodstain, marked: "Lname, FName" / BT Female 1

One approx. 2.5 x 4 inch cloth swatch with a "circle" of bloodstain, marked: "Lname, FName" / BT Male 1

BE SURE TO DISCARD ALL PACKAGING USED BY THE PT MANUFACTURER, AND AS NECESSARY, RE-PACKAGE ITEMS IN XXXXXXXXXX PACKAGING

Any questions about the specimens, call xxxxxxxxxxx or Bob Gaensslen at 312-996-2250.

Please make copies of all paperwork prepared in connection with the submission of this "case" to xxxxxx. If xxxxxx should call with any questions about the "case", please make a note of their inquiry for the files.
Send copies of all the paperwork to: Dr. R.E. Gaensslen, Forensic Science (M/C 866), UIC College of Pharmacy, 833 S. Wood Street, Chicago IL 60612-7231.

Please notify Bob Gaensslen at 312-996-2250 or Joe Peterson at 312-413-0439 immediately when results / report are received from xxxxxx.

We have agreed with participating labs to come clean about a blind PT if they suspect a case they have is a test, and if they ask us. If this "case" should be detected, and thus "blown" for some reason, we'll let you know right away.

H. REFERENCE AND TARGET LABORATORY DNA TYPING DATA

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 1-5, RFLP and PCR - 1

Male 1 / Female 1
Tests 1 - 5 RFLP PCR

	D1S7	D2S44	D4S139	D5S110	D10S28	D17S26	D17S79	DQA1	LDLR	GYPA	HBGG	D7S8	GC	D1S80
Reference 1														
BT Male 1	5332	1934	9159	5371	1483	5165								
blood	3502	1816	5460	3460	619	3085		1.3,4.1	B	AB	AB	A	C	24
M frac v-swab	5335	1943	9251	5408	1453	5166								
	3505	1813	5472	3466	615	3103		1.3,4.1	B	AB	AB	A	C	24
BT Female 1	4344	3172	8773	2604	2138	10939								
blood		3072	7267	1667	1825	4984		1.1,1.2	AB	AB	A	AB	AB	15,24
F frac v-swab	4371	3175	8842	2618	2194	10998								
		3076	7297	1657	1821	5010		1.1,1.2	AB	AB	A	AB	AB(C)	15,24
o-swab (F only)	4376	3184	8888	2618	2194	10941								
		3082	7343	1646	1822	5027		1.1,1.2	AB	AB	A	AB	AB	15,24
Reference 2														
BT Male 1	5426 / 5485	1990 / 1974	9457 / 9496	5440 / 5486	1522 / 1520		1548 / 1557	1.3,4.1	B	AB	AB	A	C	24
blood	3576 / 3580	1867 / 1864	5550 / 5551	3501 / 3516	657 / 654		1327 / 1316							
M frac v-swab	5415 / 5437	1982 / 1968	9350 / 9488	5469 / 5443	1514 / 1513		1549 / 1548	1.3,4.1	B	AB	AB	A	C	24
	3570 / 3591	1860 / 1861	5576 / 5548	3515 / 3517			1323 / 1310							

BT Female 1	4478 / 4421	3270 / 3252	9085 / 9148	2650 / 2640	2233 / 2239	1757 / 1760	1.1,1.2	AB	AB	A	AB	AB	15,24
blood		3165 / 3142	7472 / 7468	1693 / 1685	1854 / 1861	1394 / 1394							
F frac v-swab	5415 / 5345	3271 / 3252	8984 / 8981	5469 / 5454	2232 / 2248	1767 / 1762	1.1,1.2	AB	AB	A	AB	AB	15,24
	4499 / 4462	3114 / 3120	7398 / 7484	3517 / 3493	1877 / 1872	1402 / 1394							
	3599 / 3595	1975 / 1983	5578 / 5562	2650 / 2645	1515 / 1527								
		1883 / 1876		1697 / 1685									
o-swab (F only)	4415 / 4414	3238 / 3265	9049 / 9050	2650 / 2654	2232 / 2248	1760 / 1758	1.1,1.2	AB	AB	A	AB	AB	15,24
		3149 / 3108	7444 / 7443	1676 / 1685	1861 / 1865	1395 / 1394							

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 1-5, RFLP and PCR - 2

Male 1 / Female 1
Tests 1 - 5

	RFLP							PCR						
	D1S7	D2S44	D4S139	D5S110	D10S28	D17S26	D17S79	DQA1	LDLR	GYPA	HBGG	D7S8	GC	D1S80
BTest 1														
Female blood								1.1,1.2	AB	AB	A	AB	AB	
F Frac v-swab								1.1,1.2 (1.3,4.1)	AB	AB	A (B)	AB	ABC	
Male	5327	1957	9196	5415	1491			1.3,4.1	B	AB	AB	A	C	
	3514	1832	5458	3471										
M Frac v-swab								1.3,4.1	B	AB	AB	A	C	
BTest 2														
Female blood								1.1,1.2	AB	AB	A	AB	AB	15,24
F Frac v-swab								1.1,4.1*	AB	AB	A	AB	ABC	15,24
F Frac panty								1.1,1.3 4.1*	AB	AB	A*	AB	ABC*	15,24
Male								1.3,4.1	B	AB	AB	A	C	24
M Frac v-swab								1.3,4.1	B	AB	AB	A	C	24
M Frac panty								1.3,4.1	B	AB	AB	A	C	24
BTest 3														
Female Blood	4346	3171	8777	2603	2193		1711							
		3082	7271	1651	1826		1357							
Male Blood	5331	1935	9184	5374	1483		1517							
	3504	1817	5471	3460	617		1281							
Evid 3a	4333	3164	8829	2597	2191		1704							
		3076	7260	1645	1818		1354							
Evid 3b	4333	3161	8836	2595	2189		1709							
		3074	7260	1645	1818		1360							
Evid 3c	4333	3158	8789	2594	2196		1707							
		3082	7260	1645	1827		1360							
Evid 3d	4338	3172	8869	2594	2193		1707							
		3082	7269	1640	1826		1365							
Evid 3e	4342	3174	8823	2595	2193		1708							
		3074	7258	1648	1826		1358							
Evid 3f	4347	3164	8783	2597	2193		1709							
		3076	7276	1649	1825		1358							

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 1-5, STR - 1

Male 1 / Female 1
Tests 1 - 5 STR

	TPOX	THO1	CSF1PO	D3S1358	vWA	FGA	D5S818	D13S317	D7S820
Reference 1									
BT Male 1									
blood	8,9	7,9	10	17,18	19	20,24	10,13	12	10
M frac v-swab									
	8,9	7,9	10	17,18	19	20,24	10,13	12	10
BT Female 1									
blood	8	7,8	11,13	15	17	19,22	11,12	11	11
F frac v-swab									
	8(9)	7,8(9)	11,13(10)	15 (17,18)	17(19)	19,22(20,24)	11,12(10,13)	11(12)	11(10)
o-swab (F only)									
	8	7,8	11,13	15	17	19,22	11,12	11	11
Reference 2									
BT Male 1									
blood									
M frac v-swab									
BT Female 1									
blood									
F frac v-swab									
o-swab (F only)									

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 1-5, STR - 2

Male 1 / Female 1
Tests 1 - 5

	STR TPOX	THO1	CSF1PO	D3S1358	vWA	FGA	D5S818	D13S317	D7S820
BTest 1									
Female blood									
F Frac v-swab									
Male									
M Frac v-swab									
BTest 2									
Female blood	8	7,8	7,9						
F Frac v-swab	8*	7,8*	6*,7,9						
F Frac panty	8,9	7,8 9*	6*,7,9						
Male	8,9	7,9	6						
M Frac v-swab	8,9	7,9	6						
M Frac panty	8,9	7,9	6						
BTest 3									
Female Blood									
Male Blood									
Evid 3a									
Evid 3b									
Evid 3c									
Evid 3d									
Evid 3e									
Evid 3f									

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 6-10, RFLP and PCR - 1

Male 2 / Female 2 Female 3 Tests 6 - 10	RFLP						PCR						
	D1S7	D2S44	D4S139	D5S110	D10S28	D17S79	DQA1	LDLR	GYPA	HBGG	D7S8	GC	D1S80
Reference 1													
BT Male 2	7213	3383	15650	5105	4926								
blood	5165	1603	4612	2235	1863		1.2,1.3	B	AB	AB	A	AC	
M frac v-swab 2/2	7171	3364	15573	5094	4903		1.2,1.3	B	AB	AB	A	AC	
	5136	1592	4577	2221	1856								
M frac v-swab 2/3	7176	3397	16216	5120	4938		1.2,1.3	B	AB	AB	A	AC	
	5163	1602	4602	2225	1861								
BT Female 2	6760	1848	8762	5494	1835								
blood	4675	1602	6705	2955	1733		1.3,4.1	AB	A	A	AB	BC	
F frac v-swab 2/2	6730	1841	8646	5484	1830		1.3,4.1	AB	A	A	AB	BC	
	4658	1596	6664	2931	1731								
BT Female 3	9109	2668	6812	5447	1224								
blood	6105	1787	5952	1726	1128		1.2,3	AB	A	B	AB	AB	
F frac v-swab 2/3	9071	2650	6726	5406	1216		1.2,3	AB	A	B	AB	AB	
	6086	1778	5881	1702	1115								
Reference 2													
BT Male 2	7262	3429	17185	5164	5017	1787							
blood	5227	1631	4679	2254	1892	1588	1.2,1.3	B	AB	AB	A	AC	18,24
M frac v-swab 2/2	7344	3434	17328	5173	5022	1797	1.2,1.3	B	AB	AB	A	AC	18,24
	5233	1642	4643	2265	1905	1601							
M frac v-swab 2/3	7326	3447	17195	5181	5016	1789	1.2,1.3	B	AB	AB	A	AC	18,24
	5255	1639	4682	2264	1890	1589							

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 6-10, RFLP and PCR - 2

Male 2 / Female 2
Female 3 **RFLP** **PCR**
Tests 6 - 10

	D1S7	D2S44	D4S139	D5S110	D10S28	D17S79	DQA1	LDLR	GYPA	HBGG	D7S8	GC	D1S80
BT Female 2	6836	1859	8743	5525	1840	1501							
blood	4744	1619	6744	2966	1758	1299	1.3,4.1	AB	A	A	AB	BC	24
F frac v-swab 2/2	6845	1866	8766	5533	1847	1504	1.3,4.1	AB	A	A	AB	BC	24
	4722	1626	6795	2968	1763	1309							
BT Female 3	9354	2714	6895	5486	1247	1588							
blood	6237	1824	6028	1714	1141	1313	1.2,3	AB	A	B	AB	AB	24,30
F frac v-swab 2/3	9324	2685	6906	5483	1232	1574	1.2,3	AB	A	B	AB	AB	24,30
	6202	1809	6043	1728	1134	1305							
BTest 6													
Female 3 Blood							1.2,3	AB	AA	BB	AB	AB	
Male 2 buccal							1.2,1.3	BB	AB	AB	AA	AC	
F frac v-swab 2/3													
M frac v-swab 2/3							1.2,1.3	BB	AB	AB	AA	AC	

BTest 7 No band size data

BTest 8 No band size data

BTest 9 No band size data

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 6-10, STR - 1

Male 2 / Female 2
Female 3
Tests 6 - 10

	STR									
	TPOX	THO1	CSF1PO	D3S1358	vWA	FGA	D3S1358	D5S818	D13S317	D7S820
Reference 1										
BT Male 2										
blood	8,8	9.3,9.3	10,10	15,17	15,17	20,25	15,17	11,12	8,11	9,13
M frac v-swab 2/2	8,8	9.3,9.3	10,10	15,17	15,17	20,25	15,17	11,12	8,11	9,13
M frac v-swab 2/3	8,8	9.3,9.3	10,10	15,17	15,17	20,25	15,17	11,12	8,11	9,13
BT Female 2										
blood	8,11	6,9.3	11,12	16,17	15,17	18,20	16,17	11,14	11,11	10,12
F frac v-swab 2/2	8,11	6,9.3	11,12	16,17	15,17	18,20	16,17	11,14	11,11	10,12
BT Female 3										
blood	8,11	9,9	10,11	17,18	16,18	23,23	17,18	11,12	10,12	10,11
F frac v-swab 2/3	8,11	9,9	10,11	17,18	16,18	23,23	17,18	11,12	10,12	10,11
Reference 2										
BT Male 2										
blood										
M frac v-swab 2/2										
M frac v-swab 2/3										

Results of Blind Proficiency Feasibility Trials – Phase 1, Tests 6-10, STR - 2

Male 2 / Female 2
Female 3 **STR**
Tests 6 - 10

	TPOX	THO1	CSF1PO	D3S1358	vWA	FGA	D3S1358	D5S818	D13S317	D7S820
BT Female 2										
blood										
F frac v-swab 2/2										
BT Female 3										
blood										
F frac v-swab 2/3										
BTest 6										
Female 3 Blood										
Male 2 buccal										
F frac v-swab 2/3										
M frac v-swab 2/3										
BTest 7										
BTest 8										
BTest 9										

Results of Blind Proficiency Feasibility Trials – Phase 2, RFLP and PCR - 1

Phase 2 Tests Tests 11 - 15 M1, M2 & F1	RFLP						PCR						
	D1S7	D2S44	D4S139	D5S110	D10S28	D17S79	DQA1	LDLR	GYPA	HBGG	D7S8	GC	D1S80
Reference 1													
Male 1													
Male 2													
Female 1													
BT2-M2 > BT2-M1													
Q2 unspecified (is M2 victim)													
Reference 2													
Male 1							1.1,1.2	AB	A	A	A	AB	
Male 2							1.2,4.1	B	A	A	AB	BC	
Female 1							1.2,4.1	AB	A	B	B	AC	
BT2-M2 > BT2-M1													
Stains 1,2,3,5							1.2,4.1	B	A	A	AB	BC	
Stain 6							1.1,1.2	AB	A	A	A	AB	
Stain 4							1.2,4.1 (1.3)	B	A(B)	A(B)	AB	BC	
Stain 7							1.3,4.1(1.1;1.2)	B(A)	A(B)	AC(B)	AB	(A)BC	
Reference 3													
Male 1													
Male 2													
Female 1													
BT2-F1 > BT2-M1													
A Right Hip													
B Front of Right Leg													
C Front of Left Leg													
Reference 4													
Male 1							1.1,1.2	AB	A	A	A	AB	22,25
Male 2							1.2,4.1	B	A	A	AB	BC	18,24
Female 1							1.2,4.1	AB	A	B	B	AC	25,28
BT2-F1 > BT2-M1													
Pants A rt thigh (M1)							1.1,1.2	AB	A	A	A	AB	22,25
Pants B rt pocket (F1)							1.2,4.1	AB	A	B	B	AC	25,28
Pants C left thigh (F1)													

Reference 6	D1S7	D2S44	D4S139	D5S110	D10S28	D17S79	DQA1	LDLR	GYPA	HBGG	D7S8	GC	D1S80
Male 1	2461/1854	2884/2322	3747/3161	2710/1933	2967/1261	1528/1452	1.1,1.2	AB	A	A	A	AB	
Male 2	4405/3824	3533/2028	6132/4678	2849/2667	3487	1717/1528	1.2,4.1	B	A	A	AB	BC	
Female 1	6305	4114/3417	22140/8122	2581/2511	1763	1760/1375	1.2,4.1	AB	A	B	B	AC	
BT2-F1 > BT2-M1													
Q1A right upper thigh / pocket (F1)	6285	4100/3398	degr/8101	2572/2496	1759	1754/1371	1.2,4.1	AB	A	B	B	AC	
Q1B right thigh (F1)	6297	4112/3407	22006/8120	2577/2495	1758	1759/1372	1.2,4.1	AB	A	B	B	AC	
Q1C right thigh (M1)	2457/1847	2877/2314	3733/3156	2703/1937	2961/1258	1528/1449	1.1,1.2	AB	A	A	A	AB	
Q1D right thigh (F1)	6302	4106/3402	22204/8119	2576/2502	1754	1755/1373	1.2,4.1	AB	A	B	B	AC	
Q1E inside rt pants cuff (F1)	6307	4112/3398	20564/8119	2572/2493	1756	1758/1377	1.2,4.1	AB	A	B	B	AC	
Q1F left thigh (F1)	6312	4106/3407	21354/8119	2574/2501	1754	1753/1370	1.2,4.1	AB	A	B	B	AC	

Results of Blind Proficiency Feasibility Trials – Phase 2, RFLP and PCR - 2

Phase 2 Tests												
Tests 11 - 15	RFLP						PCR					
M1, M2 & F1		D2S44	D4S139	D5S110	D10S28	D17S79	DQA1	LDLR	GYPA	HBGG	D7S8	GC	D1S80
BT 11													
R Davis (v) [Male 2]							1.2,4.1	B	A	A	AB	BC	
K. Harris (s) [Male 1]							1.1,1.2	AB	A	A	A	AB	
BT2-M2 > BT2-M1													
Q1 Pants A,C,D,E,G,H							1.2,4.1	B	A	A	AB	BC	
Q1 Pants B,F							1.1,1.2	AB	A	A	A	AB	
BT 12													
Jane Morso (v) [Female 1] 2LB-1													
Alan Bruever (s) [Male 1] 3-LB1													
BT2-F1 > BT2-M1													
Pants 1-R1													
Pants 1-R2													
Pants 1-R3													
Pants 1-R4													
1-R1-1-R3 could be 2-LB1 not 3-LB1													
1-R4 could be 3-LB1 not 2-LB1													
BT 13													
Debbie Herrera (v) [Female 1]													
Ron Herrera (s) [Male 1]													
BT2-F1 > BT2-M1													
Pants (lab interprtd as Ron's)													
Q1-1													
Q1-3 and Q1-8													
BT 14													
Theresa Hunt (v) [Female 1]							1.2,4.1	AB	A	B	B	AC	
Jerome Fredericks (s) [Male 1]							1.1,1.2	AB	A	A	A	AB	
BT2-F1 > BT2-M1													
Pants most stains (1-13 and 15)							1.2,4.1	AB	A	B	B	AC	
Pants one stain (14)							1.1,1.2	AB	A	A	A	AB	
BT 15													
Anne Marie Crawford (v) [Female 1]=2A													
Ryan Allen Davis (s) [Male 1]=3A													
BT2-F1 > BT2-M1													
Pants=1 - Four areas													
area 1							S not V						
other three areas							V not S						
* DQ, PM, D1 loci - types not reported													

Results of Blind Proficiency Feasibility Trials – Phase 2, STR - 1

Phase 2 Tests Tests 11 - 15 M1, M2 & F1	STR TPOX	THO1	CSF1PO	vWA	D16S539	D7S820	D13S317	D5S818	FGA	D3S1358	D8S1179	D21S11	D18S51	Amelo
Reference 1														
Male 1	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13	19,23	17,18	12,15	28,29	17,18	
Male 2	8	9,9.3	11	17	12,13	10,12	11,12	9,10	21,24	15,17	13,14	30	17,18	
Female 1	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	
BT2-M2 > BT2-M1														
Q2 unspecified (is M2 victim)	8	9,9.3	11	17	12,13	10,12	11,12	9,10	21,24	15,17	13,14	30	17,18	
Reference 2														
Male 1	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13	19,23	17,18	12,15	28,29	17,18	XY
Male 2	8	9,9.3	11	17	12,13	10,12	11,12	9,10	21,24	15,17	13,14	30	17,18	XY
Female 1	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	XY
BT2-M2 > BT2-M1														
Stains 1,2,3,5	8	9,9.3	11	17	12,13	10,12	11,12	9,10	21,24	15,17	13,14	30	17,18	XY
Stain 6	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13	19,23	17,18	12,15	28,29	17,18	XY
Stain 4	8	9,9.3	(10),11	17(18)	(10),12,13	(8),10,12	11,12	9,10,(11),(12)	21,(23),24	15,(16),17	13,14,(15)	(28),30	17,18	XY
Stain 7	8	9,9.3	11	17	12,13	10,12	11,12	9,10	21,24	15,17	13,14	30	17,18	XY
Reference 3														
Male 1	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13						
Male 2	8	9,9.3	11	17	12,13	10,12	11,12	9,10						
Female 1	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13						
BT2-F1 > BT2-M1														
A Right Hip	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13						
B Front of Right Leg	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13						
C Front of Left Leg	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13						
Reference 4														
Male 1	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13						
Male 2	8	9,9.3	11	17	12,13	10,12	11,12	9,10						
Female 1	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13						
BT2-F1 > BT2-M1														
Pants A rt thigh (M1)	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13						
Pants B rt pocket (F1)	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13						
Pants C left thigh (F1)	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13						

	TPOX	THO1	CSF1PO	vWA	D16S539	D7S820	D13S317	D5S818	FGA	D3S1358	D8S1179	D21S11	D18S51	Amelo
Reference 6														
Male 1	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13	19,23	17,18	12,15	28,29	17,18	XY
Male 2	8	9,9.3	11	17	12,13	10,12	11,12	9,10	21,24	15,17	13,14	30	17,18	XY
Female 1	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
BT2-F1 > BT2-M1														
Q1A right upper thigh / pocket (F1)	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Q1B right thigh (F1)	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Q1C right thigh (M1)	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13	19,23	17,18	12,15	28,29	17,18	XY
Q1D right thigh (F1)	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Q1E inside rt pants cuff (F1)	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Q1F left thigh (F1)	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X

Results of Blind Proficiency Feasibility Trials – Phase 2, STR - 2

Phase 2 Tests Tests 11 - 15 M1, M2 & F1	STR TPOX	THO1	CSF1PO	vWA	D16S539	D7S820	D13S317	D5S818	FGA	D3S1358	D8S1179	D21S11	D18S51	Amelo
BT 11														
R Davis (v) [Male 2]														
K. Harris (s) [Male 1]														
BT2-M2 > BT2-M1														
Q1 Pants A,C,D,E,G,H														
Q1 Pants B,F														
BT 12														
Jane Morso (v) [Female 1] 2LB-1	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Alan Bruever (s) [Male 1] 3-LB1	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13	19,23	17,18	12,15	28,29	17,18	X,Y
BT2-F1 > BT2-M1														
Pants 1-R1	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Pants 1-R2	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Pants 1-R3	11	9.3	10,12	19,20	10,13	10,11	8,11	11,13	20,24	15,17	13,16	28	12,17	X
Pants 1-R4	8,11	7,9.3	10,11	14,16	10,11	8,10	12	9,13	19,23	17,18	12,15	28,29	17,18	X,Y
1-R1-1-R3 could be 2-LB1 not 3-LB1														
1-R4 could be 3-LB1 not 2-LB1														
BT 13														
Debbie Herrera (v) [Female 1]														
Ron Herrera (s) [Male 1]														
BT2-F1 > BT2-M1														
Pants (lab interprtd as Ron's)														
Q1-1	S not V			x		x	x	x	x	x	x	x	x	
Q1-3 and Q1-8	V not S													
BT 14														
Theresa Hunt (v) [Female 1]														
Jerome Fredericks (s) [Male 1]														
BT2-F1 > BT2-M1														
Pants most stains (1-13 and 15)														
Pants one stain (14)														
BT 15														
Anne Marie Crawford (v) [Female 1]=2A														
Ryan Allen Davis (s) [Male 1]=3A														
BT2-F1 > BT2-M1														
Pants=1 - Four areas														
area 1														
other three areas														

I-1. LABORATORY POST-TEST NOTIFICATION

Director
xxx DNA Laboratory
xxxxxxxxxxxxx
xxxxxxxxxxxxx

Dear xxxxx:

We write to inform you and confirm that the case represented by your Case Number xxxxx [Requesting Agency No xxxxx, xxxxx Police Dept], submitted to you by the xxxxx (in the matter of a homicide: xxxxx, victim; xxxxx, suspect; offense date 7/8/97), was a blind proficiency feasibility test case. This "case" was fictitious, the people named in the "case" do not exist, and the evidentiary items were manufactured.

We want to take this opportunity to thank you for your initial willingness to participate in this blind PT feasibility project, and for your participation in an actual trial. The biological specimens submitted in this "case" were typed in two different reference laboratories for a series of RFLP loci, HLA-DQA1 and PM loci, D1S80, and several STR loci. Attached, we share the results obtained by the reference laboratories, as well as those of the other blind trial target test labs who have returned results to date, for your information. All the data received thus far is in good agreement. The "suspect" in your "case" is "BT Male 1" on the reference sheets; the "victim" is "BT Female 1."

In accordance with our written agreement with you (and with all other participants in this feasibility study), we will not reveal in any public forum or in any public documents (nor will we reveal to the NIJ in our reporting) the identities of the participating labs nor of the participating law enforcement agencies. The extent to which you wish to reveal your own participation to external parties or to people within your own laboratory operation is, of course, up to your discretion.

The records associated with this "case" should be flagged at this time, to be handled by your laboratory record keeping system as a blind proficiency feasibility test, as distinct from a real case. If you have entered any of the DNA profiles from this "case" into any database or databank, the profile(s) should be removed at this time. To give our biological-specimen donors further confidence in our representations to them that any such profiles would be purged from databanks or databases, we would ask that you complete and sign the attached form and return it to us at your convenience.

We again thank you for your participation in the feasibility test. We expect that the findings will be carefully considered by the DAB and other policy makers in deciding on the value of blind PT in future recommended or required QA/QC programs in the nation's forensic DNA testing laboratories.

With best wishes,

R.E. Gaensslen
Joseph L. Peterson
Project Directors, NIJ Blind DNA PT Project

I-2. LAW ENFORCEMENT POST-TEST NOTIFICATION

xxxxxx
xxxxx Police Department
xxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxl

Dear xxxxx:

We write to confirm that we have formally revealed to the xxxxx Lab that Case # xxxxx (Laboratory No. xxxxx) submitted by you to them on our behalf was a blind proficiency feasibility case.

We want to take this opportunity to thank you for your willingness to participate in this blind PT feasibility project, and for your participation in one of the actual trials. The biological specimens submitted in this "case" have been or will be typed in five different reference laboratories, and we will share the results obtained by those labs, as well as the results obtained by other blind-tested labs, with the laboratory director. The lab's conclusions in this case are completely concordant with the way the evidence item was manufactured.

In accordance with our written agreement with you (and with all other participants in this feasibility study), we will not reveal in any public forum, in any public documents, or to the governmental agency sponsor, the identities of the participating labs nor of the participating law enforcement agencies. The extent to which you wish to reveal your own participation to external parties or to people within your own operation is, of course, up to your discretion.

The records associated with this "case" could be flagged at this time, if necessary, to distinguish them from real case records in the departmental record keeping system.

We again thank you for all your cooperation in helping us construct and submit the feasibility test. We expect that the findings will be considered by the DNA Advisory Board and other policy makers in deciding on the value of blind PT in future recommended or required QA/QC programs in the nation's forensic DNA testing laboratories.

With best wishes,

R.E. Gaensslen
Joseph L. Peterson
Project Directors, NIJ Blind DNA PT Project

c: xxxxx, Chief of Police

I-3. LABORATORY POST-TEST CODIS PURGE CERTIFICATION

University of Illinois at Chicago
NIJ Blind DNA Proficiency Testing Study

Certification by Blind-Tested Laboratory of Removal of DNA Profiles from Computer Databases / Databanks

Name / Address of Laboratory:

xxxxxx DNA Laboratory
xxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxx

It has been confirmed in writing to us that our Case Number xxxxx, Xref xxxxx [Submitting Agency, xxxxx, Number xxxxx] submitted by the xxxxx, in the matter of a homicide (xxxxx, victim; xxxxx, suspect; offense date 7/8/97) was a blind proficiency feasibility test, that this case was not real, and that the named persons in the case are fictitious.

This laboratory affirms, by signature of an authorized official, that: (i) the DNA profile(s) obtained in the above-captioned fictitious case have not been, and will not be, entered into any private, local, state, or federal DNA typing database or databank; OR that (ii) any DNA profile(s) obtained in the above-captioned fictitious case that were entered into any private, local, state, or federal DNA typing database or databank have been permanently removed.

Authorized Official of the Laboratory:

_____
Typed Name

_____
Signature

_____
Date Signed

Date received by UIC Project Office: