Automated Laboratory Standards:

RESULTS FROM THE SURVEY OF LABORATORY AUTOMATED DATA MANAGEMENT PRACTICES

Prepared for:

Office of Information Resources Management
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

June 15, 1990

Prepared by:

BOOZ•ALLEN & HAMILTON Inc.
4330 East-West Highway
Bethesda, Maryland 20814
(301) 951-2200

Contract No. 68-W9-0037

Computer Sciences Corporation
79 T.W. Alexander Drive
Research Triangle Park, North Carolina 27709
(919) 541-9287

Contract No. 68-01-7365

Acknowledgments

This report was the combined efforts of Computer Sciences Corporation, Booz•Allen & Hamilton Inc., EPA staff, and outside experts. Terrie Baker of CSC researched and prepared the draft for public review. Marguerite Jones, Ronald Ross, Lynn Eberhardt, and Linda Palumbo of Booz•Allen evaluated the comments and completed this final report. Numerous EPA staff and outside experts provided substantial critical reviews and valuable technical comments. Richard Johnson of the Scientific Systems Staff of EPA's Office of Information Resources Management directed the contractors' work and managed the review process.

Table of Contents

Executive Summary

The U.S. Environmental Protection Agency (EPA) has initiated a program to ensure the integrity of computer-resident data in laboratories performing analyses in support of EPA programs by developing standards for automated laboratory processes. The possession of sound technical data provides a fundamental resource for EPA's mission to protect the public health and environment.

This report describes the findings of a survey of laboratories engaged in analytical chemistry in support of EPA programs and employing automated information systems to generate, analyze, and report the findings. A survey questionnaire was developed from existing standards for laboratory and automated operations. Five areas of automated technology management were addressed: organization, security, documentation, operations, and traceability.

The results of the data analysis revealed that in the majority of areas surveyed, the respondent laboratories were in compliance with many of the already established Good Laboratory Practice regulations (GLPs) and standards for automation. However, four areas were identified as having substantial deficiencies with respect to meeting the standards: autonomy of the quality assurance unit, system security, system documentation, and practices for data editing.

By definition, a quality assurance unit or group must be independent of day-to-day laboratory operations to provide an unbiased review of the quality of work conducted. In many of the laboratories, the quality assurance function was not independent, in that the individual responsible for quality assurance typically reported to the laboratory chief or manager; in some of the laboratories, that individual was the laboratory manager.

System security was the most variable and showed the highest inconsistency between respondent laboratories of all the areas evaluated by the survey. When asked if the risks and associated protection requirements

were determined by a formal risk analysis, 91 percent of the respondents answered that no risk analysis had been conducted. **In more than nine out of ten cases, no specific standards or other guidance were used in the design and implementation of security measures.** Only half the laboratories reported that they train new staff in the security procedures of their data management systems, and no respondents reported that they offer periodic refresher courses to existing personnel.

None of the laboratories surveyed had a full complement of standard system documentation as specified by EPA policy. Interestingly enough, roughly 65 percent of the laboratories surveyed felt their available documentation represented a conscientious effort to document the system software. Clearly, in the area of system documentation, there exists a need to educate laboratory personnel in the steps necessary to achieve compliance with system documentation standards.

Procedures for data changes or editing were often in conflict with GLP requirements, which state that data editing must be clearly documented to include when changes were made, who made the changes, and why. More than four out of five respondents indicated that there was no written documentation requirement for who requested or who made the change to the system, and less than half reported that their systems kept a log of the data change information.

The results of this survey indicate that there is a need for standardization in the data management procedures used in analytical chemistry laboratories supporting various EPA programs. The Agency should assume responsibility for establishing standards for safeguarding the security of computer-resident laboratory data. Sound data are a fundamental resource for EPA's mission to protect the public health and the environment.

Background

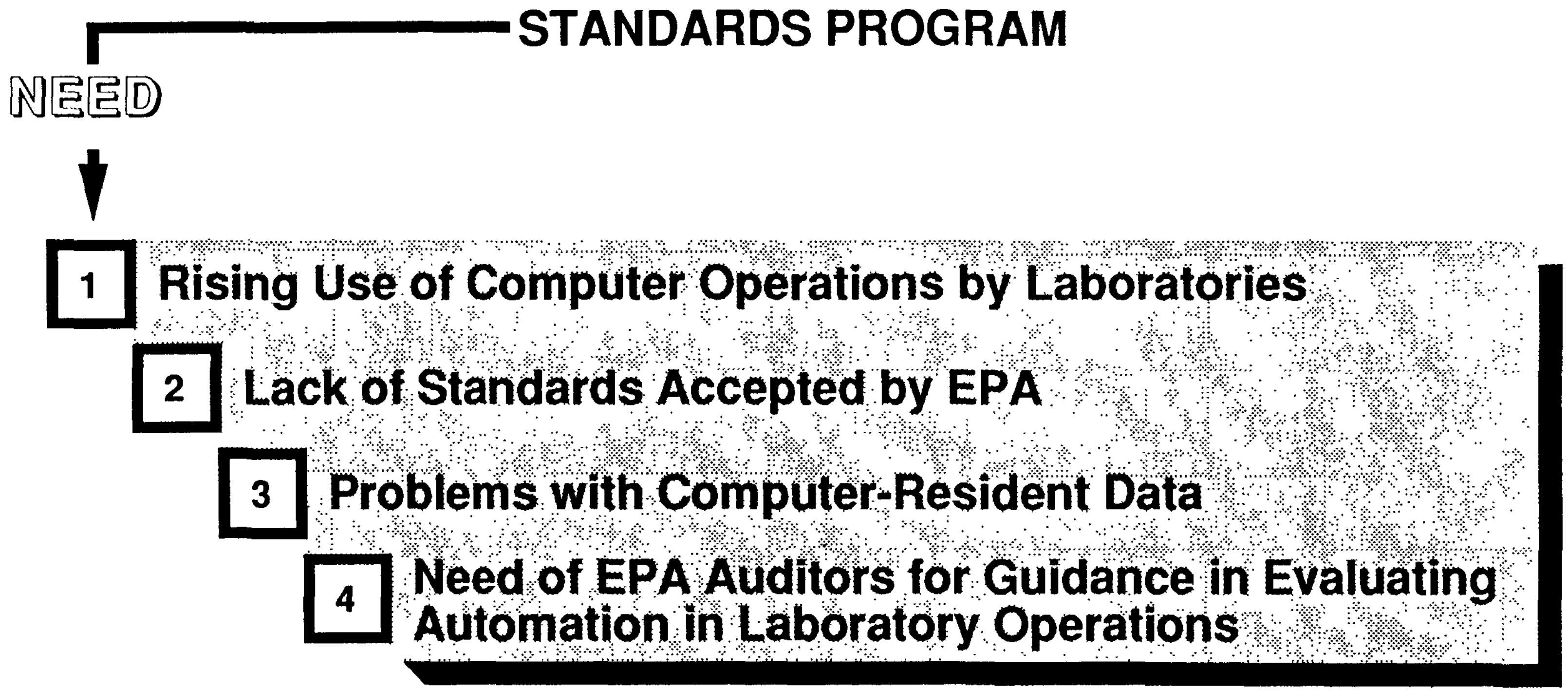
The U.S. Environmental Protection Agency (EPA) has initiated a program to ensure the integrity of computer-resident data in laboratories performing analyses in support of EPA programs by developing standards for automated laboratory processes. The possession of sound technical data provides a fundamental resource for EPA's mission to protect the public health and environment, implemented through several environmental programs. The activities of these environmental programs are diverse, and include basic research at EPA's environmental research centers, environmental sample analyses at EPA's regional laboratories and contractors' laboratories, and product registration relying on analytical data submitted by the private sector.

EPA recognizes that the implementation of an automated laboratory standards program will require each laboratory to allocate resources of dollars and time for the program's execution. Experience has shown that in developing and using a proper standards program, a net savings may be achieved, as acquisition, recording, and archiving of data will be improved with a net reduction in test duplication.

Within EPA, the Office of Information Resources Management (OIRM) has assumed the objective of establishing an automated laboratory standards program. The need for this program is evidenced by several factors. Exhibit 1 illustrates these factors, which include the rising use of computerized operations by laboratories, the lack of uniform standards developed or accepted by EPA, evidence of problems associated with computer-resident data, and the evolving needs of EPA auditors and inspectors for guidance in evaluating automated laboratory operations.

Laboratories collecting data for EPA's programs have taken advantage of increasing technology to streamline the analytical processes. Initially, automated instrumentation entered the laboratories to increase productivity and enhance the accuracy of reported results. Then, computers maintaining data bases of results were used for data management and tracking. These

EXHIBIT 1
Need for EPA's Automated Laboratory Standards Program



STANDARDS PROGRAM

NEED

1 Rising Use of Computer Operations by Laboratories

2 Lack of Standards Accepted by EPA

3 Problems with Computer-Resident Data

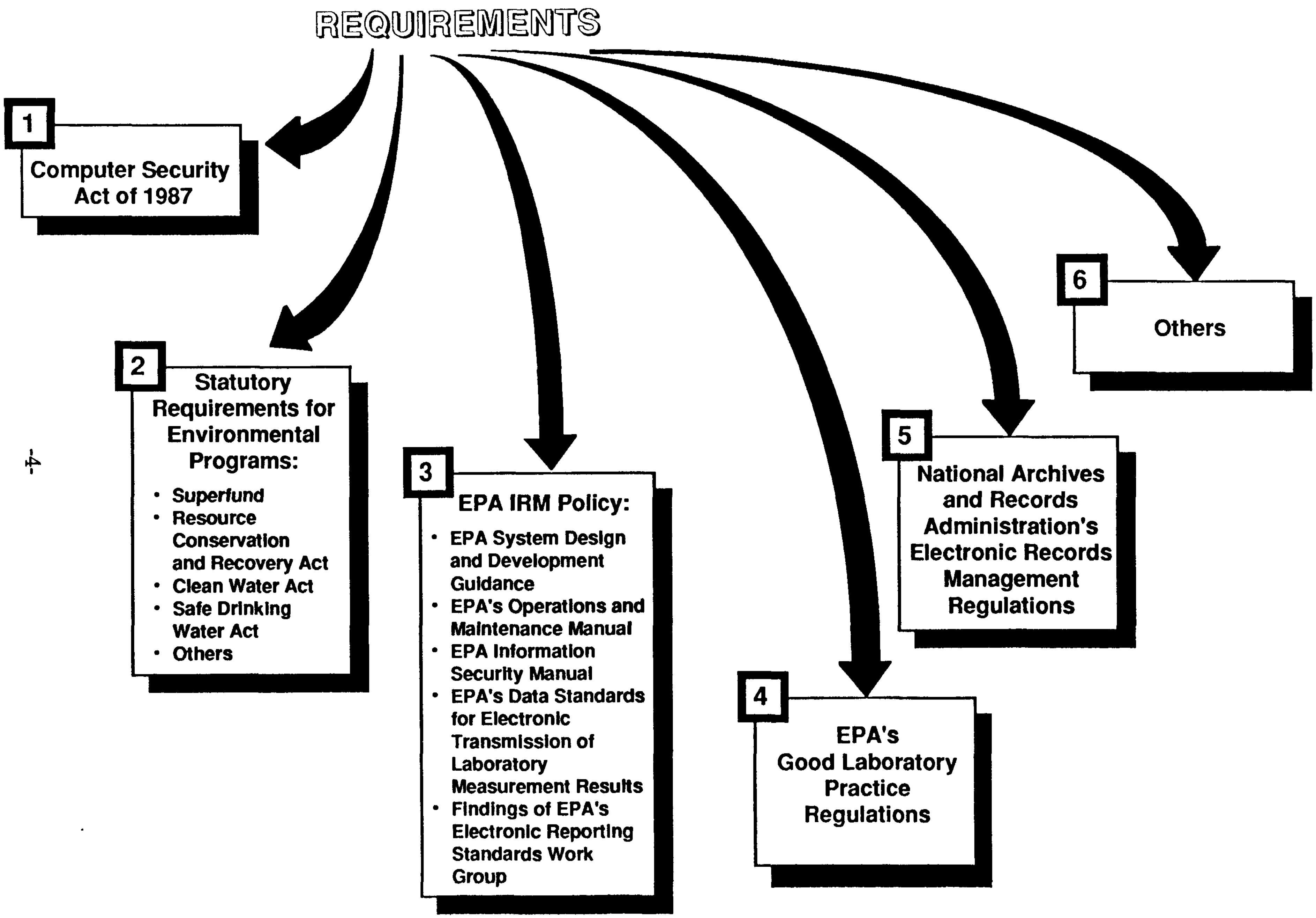
4 Need of EPA Auditors for Guidance in Evaluating Automation in Laboratory Operations

computer systems were integrated into more sophisticated laboratory information management systems (LIMS). Methods for data reporting include electronic mail, electronic bulletin boards, and direct links between central processing units. Each of these advances necessitates thorough quality control procedures for data generation, storage, and retrieval to ensure the integrity of computer-resident data.

Currently, EPA has no Agency-wide principles that laboratories collecting and evaluating computer-resident data must follow. The requirements that must be considered in developing automated laboratory standards come from a variety of sources, as Exhibit 2 illustrates, including the requirement of the Computer Security Act of 1987 (P.L. 100-235, January 8, 1988) and various EPA program-specific data collection requirements under Superfund, the Resource Conservation and Recovery Act, the Clean Water Act, and the Safe Drinking Water Act, among others. Additionally, OIRM has developed electronic transmission standards and is developing a strategy for electronic record keeping and electronic reporting standards that will impact on all Agency activities. The development of uniform principles for automated data in EPA laboratories, regardless of program, will take into account the common elements of all these data collection activities, and provide a minimum standard that each laboratory should achieve.

There is increasing evidence of problems associated with the collection and use of computer-resident laboratory data supporting various EPA programs. To illustrate, as of November 1989, EPA's Office of the Inspector General was investigating between 10 and 12 laboratories in Superfund's Contract Laboratory Program (CLP) for a variety of allegations, including "time traveling" and instrument calibration violations. In "time traveling," sample testing dates are manipulated, by either adjusting the internal clock of the instrumentation performing the analyses or manipulating the resultant computer-resident data. (Hazardous waste samples must be assayed within a prescribed time period or the results may be compromised.) Additionally, calibration standard results have allegedly been electronically manipulated and other calibration results substituted when the actual results did not meet the range specifications of the CLP procedure being followed. If proven, these allegations may be treated as felonies.

EXHIBIT 2
Considerations in Developing Automated Laboratory Standards



REQUIREMENTS

1 Computer Security Act of 1987

2 Statutory Requirements for Environmental Programs:
- Superfund
- Resource Conservation and Recovery Act
- Clean Water Act
- Safe Drinking Water Act
- Others

3 EPA IRM Policy:
- EPA System Design and Development Guidance
- EPA's Operations and Maintenance Manual
- EPA Information Security Manual
- EPA's Data Standards for Electronic Transmission of Laboratory Measurement Results
- Findings of EPA's Electronic Reporting Standards Work Group

4 EPA's Good Laboratory Practice Regulations

5 National Archives and Records Administration's Electronic Records Management Regulations

6 Others

Because the introduction of automation is relatively new and still evolving, no definitive guidelines for EPA auditors and inspectors have been developed. Inspectors must be alert to those steps in the procedures used by laboratories generating and using computer-resident data where the greatest risks exist. These critical process points indicate the magnitude of control that should be placed on each step of the process. If adequate controls are not present, the remainder of the process cannot correct a deviation, and the entire process will provide no reliable conclusions. Automation introduces many new variables into a system, each with its own set of critical process points. Inspectors must verify that laboratory management has recognized the various risks and has instituted an appropriate risk management program.

As part of EPA's program to ensure the integrity of automated laboratory data, OIRM developed a survey to collect information on current automated technology management practices at contract laboratories and EPA regional laboratories. The purpose of the survey was to obtain a detailed picture of the laboratory management practices. Other areas of evaluation in developing the standards program include a survey of current automated technology, a review of EPA's Good Laboratory Practice (GLP) standards developed for the pesticide and the industrial chemical testing programs (U.S. EPA, 1989a; U.S. EPA, 1989b), and an evaluation of the use of automated financial system procedures. The findings of each of these evaluations are provided in separate reports.

This document presents the findings of the automated technology practices survey. It is intended to provide guidance in developing and implementing standards for automated laboratories by evaluating existing procedures and determining any inconsistencies with EPA's GLP regulations and other current standards.

Procedures

A survey questionnaire (OIRM, 1989a) was developed from existing standards for automated operations including the Computer Security Act, the Agency's Data Standards for the Electronic Transmission of Laboratory Measurement Results, the EPA System Design and Development Guidance, and from the GLPs promulgated by the Office of Toxic Substances (40 CFR Part 792) and the Office of Pesticides Programs (40 CFR Part 160). The survey form was designed to be completed by the quality assurance manager or someone not directly responsible for laboratory and/or systems operations. Five areas of automated technology management were addressed: organization (including the system and personnel), security, documentation, operations, and traceability.

The survey form was mailed to 140 Superfund CLP laboratories and the 10 EPA regional laboratories. Eighteen laboratories (4 regional and 14 CLP laboratories) responded, providing information about 25 computer systems. Data from the completed questionnaires were entered into the Macintosh-based spreadsheet, Microsoft Excel. Responses to each question were then tabulated and analyzed. Averages were calculated for numerical responses. Complete quantitative findings are included in Appendix A to this report. Appendix A also shows the response rate to each question. The low response rate to many of the questions makes it difficult to perform a thorough statistical analysis on the findings. Appendix B to the report shows qualitative information, including a description of the systems available to each of the laboratories, and the key personnel for the laboratories. Thus, only those areas where trends were apparent are discussed.

Findings

The response rate in the survey was approximately 10 percent for Superfund CLP laboratories and 40 percent for EPA regional laboratories. The quantitative and qualitative findings presented below and in Appendices A and B are intended to provide overall trends on the state of laboratory management procedures in many of the laboratories generating data in support of EPA programs. These trends may not be consistent across all laboratories generating data for each EPA program. The findings below are presented in the order in which they appear in the survey document.

ORGANIZATION

System Organization

System Identification

The complexity, features, and constraints in operating a data management system depend heavily on the components selected. The data management systems used by the respondent laboratories are widely varied, and a matrix of the laboratories and their system descriptions can be found in Appendix B.

System Environment

When storing mission-critical data, use of an uninterruptible power supply (UPS), such as a battery or generator, is of the utmost importance. The purpose of the UPS is to make the data management system independent of the electrical grid in case of power surges, spikes, or brown-outs. A system that does not use the safety advantages of a UPS is susceptible to a breech in data integrity in cases of power interruption. Three quarters of the respondent laboratories do not operate their systems in conjunction with a UPS. It was reported that there is an average of approximately 36 hours of system down time due to power outages per year.

In general, most computer hardware was installed to minimize environmental hazards, such as exposure to moisture, temperature extremes, electrical surges, and corrosive atmospheric conditions.

Personnel

Responsible Person

The length of experience of the person in charge of the data management systems in the laboratory is an important factor and directly relates to the ability to recognize and alleviate common system problems. The survey data revealed a range in the years of experience of more than 25 years, with the average being seven years.

Most respondents have duties other than those pertaining to the data management systems they oversee (see Appendix B). On average, these responsible persons spend only a third of their time on data management responsibilities. This can explain the finding that most laboratories reported that neither data processing personnel, their supervisors, nor personnel pertaining to the system were available at all times to answer questions.

Quality Assurance Personnel

If individuals responsible for quality assurance in the laboratories report to someone directly responsible for laboratory operations, they may be susceptible to questions of conflict of interest. In most cases, the quality assurance officer was found to report to the laboratory chief or director. Additionally, quality assurance personnel should not have direct responsibility for the day-to-day operations of the laboratory to avoid any conflict-of-interest questions. In this survey, the individual that supervises the quality assurance group in most cases did not have direct responsibility for the day-to-day laboratory operation for the EPA work. However, the laboratory chief was the quality assurance officer in a few cases (see Appendix B).

Quality assurance personnel and procedures are critical to the mission of an automated laboratory data management system. The integrity of the data management system is a high priority, and the data integrity of the system is directly related to the experience and talent of the quality assurance personnel. Most of the quality assurance personnel with responsibilities relating to the surveyed laboratories did not use a data management system on a daily basis.

Staff Training and Experience

The GLPs state that each individual engaged in the conduct of or responsible for the supervision of laboratory studies must have the education, training, and experience to enable that individual to perform the assigned functions [40 CFR 792.29(a)]. Overall, the training practices and task assignment procedures of the majority of laboratories surveyed are in line with GLP and other guidelines. The only area of deficiency identified in the training and experience of staff is three quarters of the respondents reported that there is no system in place for the documentation of the training of new staff members.

SECURITY

Security Needs and Risk Assessment

Information is an Agency asset that must be safeguarded and used efficiently, just as personnel time and funds are assets. As with other assets, information resources are exposed to potential loss and misuse from a variety of accidental and deliberate causes. The extent of the potential risk must be assessed, and the level of security needs must be addressed. Of all the areas evaluated by the survey, security possessed the highest variability and lack of continuity between respondent laboratories. This seems to suggest a need for increased standardization in this area.

System protection requirements are defined in terms of confidentiality, integrity, and availability. Data confidentiality refers to a system that contains information requiring protection from unauthorized disclosure. Data integrity concerns a system that contains information requiring protection from unauthorized modification. Data availability concerns a system containing information or providing services that must be available on a timely basis to meet mission requirements.

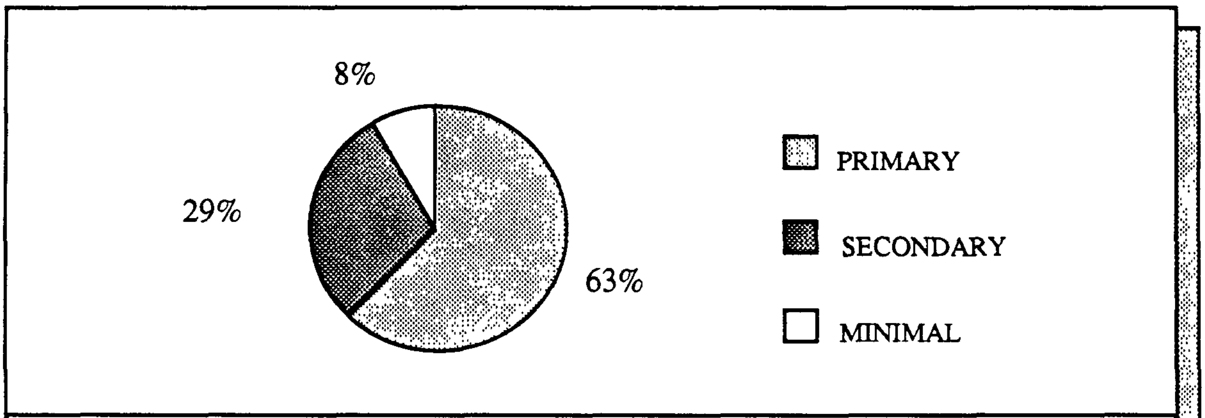
The confidentiality, integrity, and availability of the data resident in automated management systems should be maximized to best meet the needs of the users and ensure accuracy of the data within. Exhibits 3, 4 and 5 on the following pages illustrate whether the respondents' need for data confidentiality, integrity, or availability is primary, secondary, or minimal. Only two respondents indicated a minimal need for confidentiality, and no one indicated a minimal need in the other two areas. These results indicate that these three areas are of high priority in most laboratory environments.

When asked if the risks and associated protection requirements were determined by a formal risk analysis, 91 percent of the respondents answered that no risk analysis had been conducted. In more than nine out of ten cases, no specific standards or other guidance were used in the design and implementation of security measures. Another concern in the area of security is the training of new personnel. Only half the laboratories surveyed reported that they train new staff in the security procedures of their data management systems, and no respondents reported that they offer periodic refresher courses on security topics to existing personnel.

System Access Security

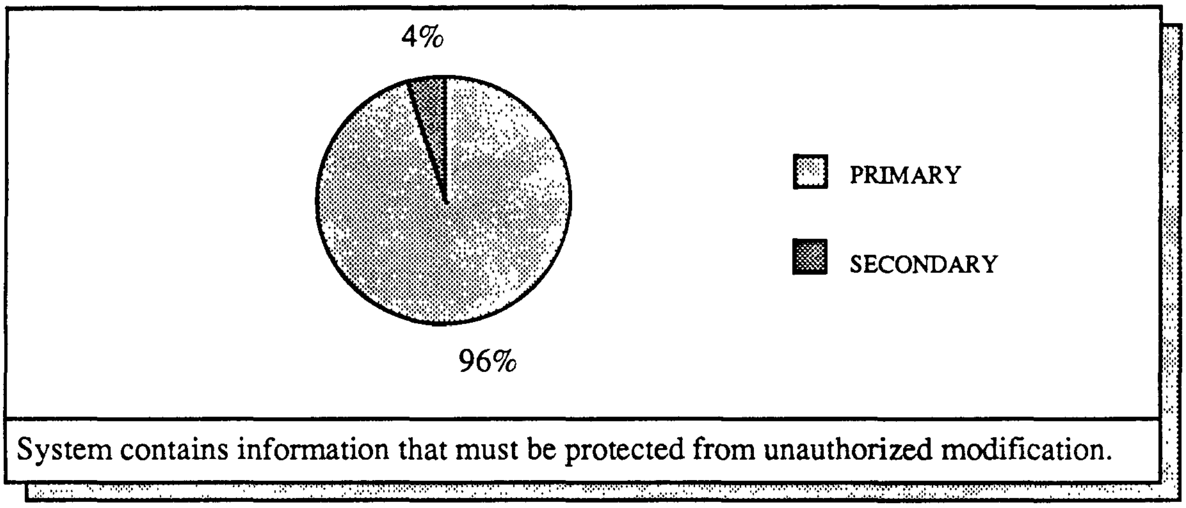
Implementing system access security standards is a proven alternative if there is a need to safeguard data input, modification, or retrieval capabilities. These standards may consist of personalized log-on requirements, individual passwords, limited access files, or data edit flags, to name a few. It was reported that 56 percent of the systems required personalized log-ons for each user, but the majority of respondents operate without an established password standard. To ensure the integrity of the information in any data management system, access security should be implemented across the board.

EXHIBIT 3
Importance Of Data Confidentiality



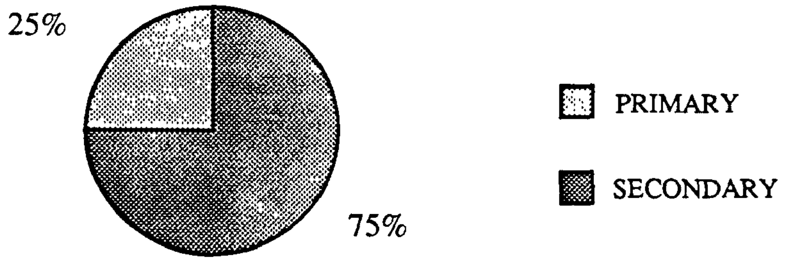
System contains information that must be protected from unauthorized disclosure.

EXHIBIT 4
Importance Of Data Integrity



System contains information that must be protected from unauthorized modification.

EXHIBIT 5
Importance of Data Availability



System contains information or provides services that must be available on a timely basis to meet mission requirements or to avoid other types of losses.

DOCUMENTATION

System Documents

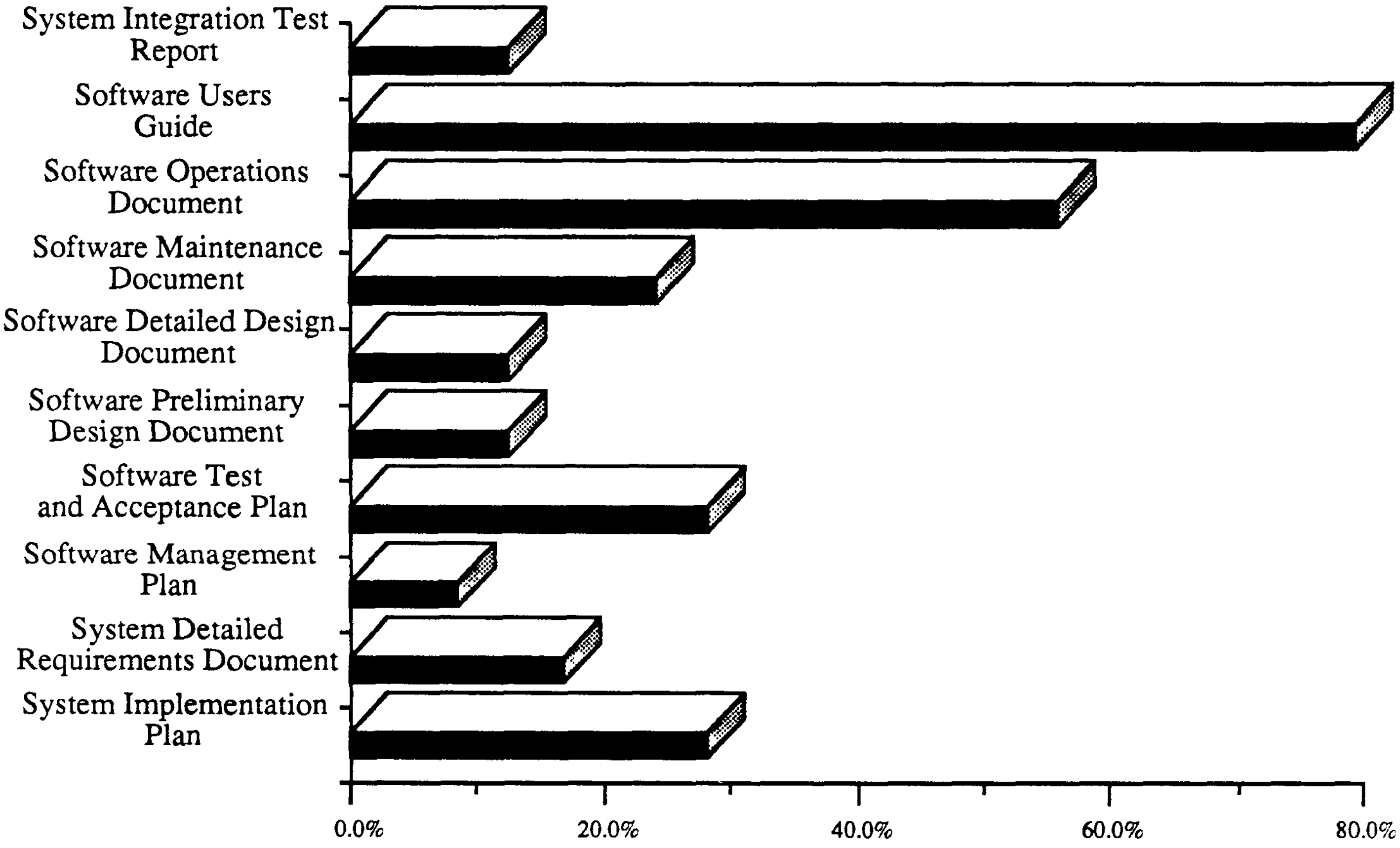
EPA policy requires that system design and development follow a typical life cycle; each step in the life cycle must be documented to demonstrate that system development efforts were conducted efficiently and effectively (OIRM, 1989b). The laboratories in the survey were asked to respond to questions regarding the extent of system documentation. Specifically, the documents in question were the following:

- System Implementation Plan
- System Detailed Requirements Document
- Software Management Plan
- Software Test and Acceptance Plan
- Software Preliminary Design Document
- Software Detailed Design Document
- Software Maintenance Document
- Software User's Guide
- System Integration Test Reports.

All the above-mentioned documents are standard requirements in the implementation of automated information systems (OIRM, 1989b). None of the laboratories surveyed had a full complement of this documentation. Exhibit 6 illustrates the percentage of laboratories having each required system document. Interestingly enough, roughly 65 percent of the laboratories surveyed felt their available documentation represented a conscientious effort to document the system software. Clearly, in the area of system documentation, there exists a need to educate laboratory personnel in the steps necessary to achieve compliance with system documentation standards.

EXHIBIT 6

Percent of Laboratories Meeting Requirements For System Documentation



OPERATIONS

Data Entry

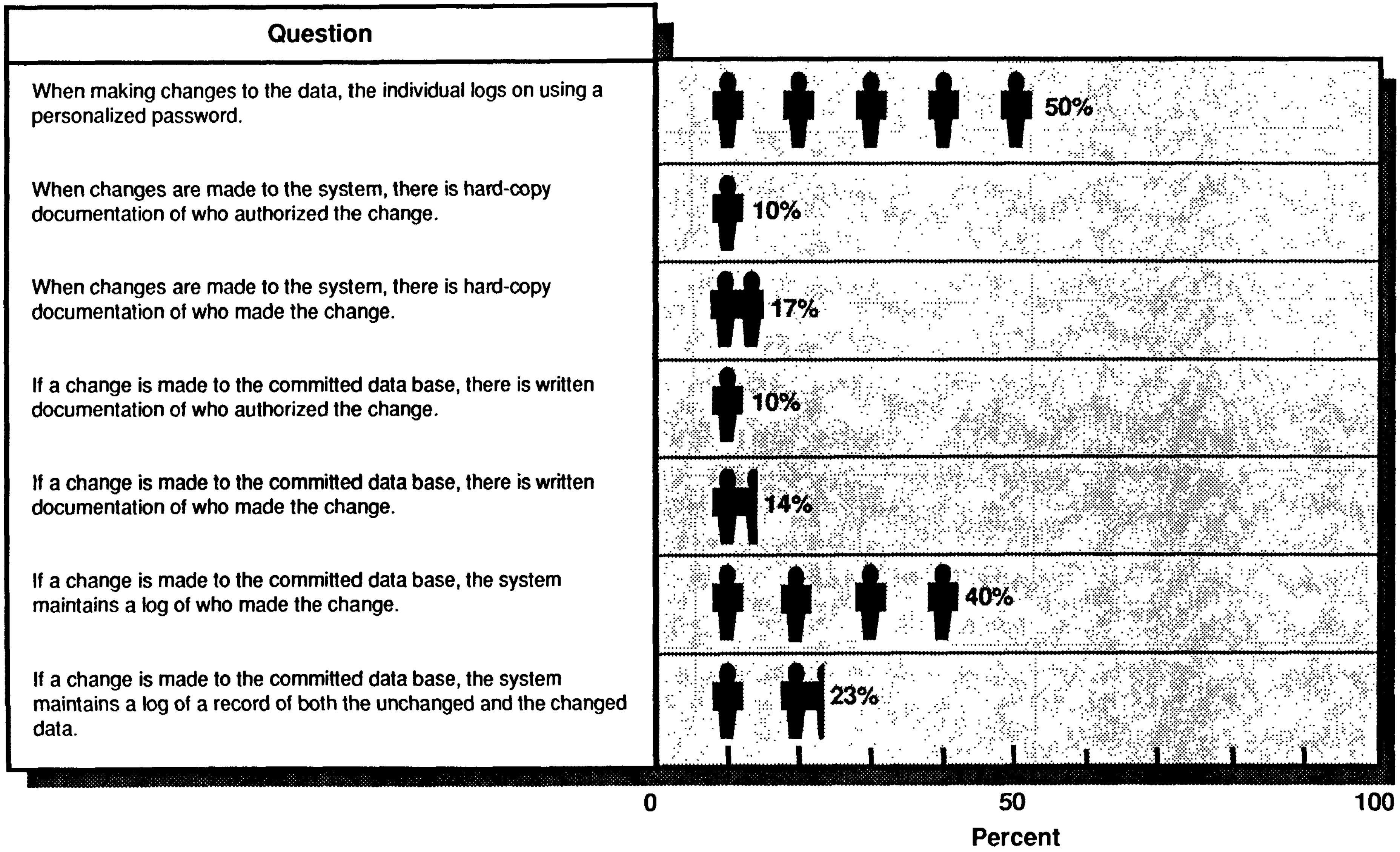
Because data integrity in an automated system is most vulnerable during entry, criteria must be established for data validation. It was found that for the laboratories that manually entered data into their systems from hard copy, **in almost every case, validation of those data is facilitated by review by the personnel that originally keyed the data rather than re-keying by another person.** Only half the laboratories responded that they use a system that prevents the entry of incorrect or out-of-range data. Seventy-one percent of the respondent laboratories reported that they do not test their new employee trainees by requiring them to demonstrate their proficiency with new assignments.

Automated laboratories may employ several data entry practices, such as personalized log-on, password requirements, various system checks, tests, and alarms if data are entered incorrectly, to minimize error and enhance data ownership during data entry. Although the implementation of the above-mentioned data entry practices could be improved, the survey findings indicate that the majority of laboratories are employing one or more of these data entry practices.

Data Changes

The GLPs state that data editing must be clearly documented to include when changes were made, who made the changes, and why. The original data entry must not be overwritten, made illegible, or deleted. Hard-copy or on-line system documentation of any changes to the data base is an essential practice to maintain the integrity of the data management system. When the data were committed to the data base or a change was made to the committed data base, more than four out of five respondents indicated that there is no written documentation requirement for who requested or who made the change to the system, and less than half reported that their systems kept a log of the data change information. Exhibit 7 gives a visual illustration of these and other findings on data change practices.

EXHIBIT 7
Extent of Use of Selected Data Change Practices



Question	
When making changes to the data, the individual logs on using a personalized password.	50%
When changes are made to the system, there is hard-copy documentation of who authorized the change.	10%
When changes are made to the system, there is hard-copy documentation of who made the change.	17%
If a change is made to the committed data base, there is written documentation of who authorized the change.	10%
If a change is made to the committed data base, there is written documentation of who made the change.	14%
If a change is made to the committed data base, the system maintains a log of who made the change.	40%
If a change is made to the committed data base, the system maintains a log of a record of both the unchanged and the changed data.	23%

Data Reduction, Analysis, and Assessment

OIRM policy states that algorithms performed by the system for data manipulation must be documented in written format (OIRM, 1989b). The majority of the respondents indicated that the algorithms were available in written format. This requirement should be continued by those who use it and adopted by those who do not. Before the algorithms are used for data analysis, they are reviewed in written format for accuracy by quality assurance staff most of the time. It was found that the algorithms are usually checked during system development, but are rarely rechecked staff on a periodic basis. Additionally, any modifications made to algorithms are documented more times than not. However, functional testing of the algorithms against test data records, although typically conducted, is not usually documented or reviewed by the quality assurance staff.

Data Outputs

The results of analyses and subsequent data reduction are used for decision-making purposes. The data reports, or outputs, containing these results must be constructed appropriately, available on a timely basis, and accurate. Most of the responding laboratories have written procedures for generating reports, graphs, and charts. Laboratory staff responsible for report generation are almost always trained. Training was typically on the job, with close instruction by supervisors, or less frequently, by co-workers.

Although the staff in half the laboratories have experienced delays in computer-generated reports that can hamper job efficiency, about 90 percent of the time, each system generates reports on a timely basis.

When final reports are generated, 90 percent of the laboratories have no typical need to manipulate the data any further. However, only about 30 percent of the laboratories "lock" the associated data base so that no further changes can be made to the data. If the data were subsequently changed, reprints of final reports may not, then, be exact duplicates of the original hard-copy report.

Back-ups/Archival

To ensure that mission-critical information is available and not lost due to power outages or flaws in magnetic media, data should be routinely backed up. Additionally, program-specific data are subject to records-retention requirements and must be archived. All the laboratories reported that they make periodic system back-ups either daily, weekly, or monthly. More than 90 percent of those back-ups performed are total system back-ups. Eight out of ten respondents said for long-term storage, the back-up media are stored off site, and the majority indicated that the media are kept in a fire proof area.

System Maintenance

As with any piece of laboratory equipment, the computer system must be maintained to meet operating specifications. This maintenance can be as repairs are required or on a regular basis. Ninety percent of the laboratories have an individual designated as responsible for system maintenance. About half the laboratories reported having a regularly scheduled preventative maintenance program, typically monthly. About two-thirds of the laboratories document the preventative maintenance program, including the length of system down time.

Repair Service

Laboratories frequently receive service contracts that offer routine or problem-solving services for computer systems purchased from vendors. In most cases, respondents have service contracts in place, and the typical response time by service technicians is approximately two hours. Most laboratories have provisions to continue laboratory operations if the system is down.

Recovery from System Failure

To preserve the integrity of computer-resident data, a disaster recovery plan must be in place to compensate for system crashes or other fault. Of the respondent laboratories, 81 percent reported they currently have no disaster recovery plan.

TRACEABILITY

Records Tracking

GLPs require that records associated with analyses, such as instrument calibrations, quality control records and other material, must be maintained in addition to the raw findings data. Respondents were asked if records of instrument calibrations, quality control samples, and data flag information were maintained in on-line or hard-copy form. One quarter of the respondents surveyed did not respond to the inquires in this section. If one assumes that the lack of a response indicates the questions is not applicable, then it is assumed that in one out of four cases among the survey respondents, records tracking standards of either type were not in place. This indicates a deficiency in this area and further study may be required.

Records Audit

Laboratories should conduct periodic self-audits to ensure that all information contained in the data system accurately reflect the raw data needed for decision-making purposes. Most of the laboratories surveyed reported that their systems were capable of supporting reports audit operations of some type. These audit functions could be linear or quadratic reduction for standard curves, quantitative analysis of unknowns, flagging of data to indicate sample results outside of some predetermined linear range, or a written record of data manipulation by the system, to mention just a few. The audit operations varied depending on the types of audit functions employed and the degree to which they were conducted. The majority of laboratories do not perform common data reduction functions such as flagging data to indicate that the standards that have been run concurrently are outside of the quality control acceptance criteria or flagging sample results to indicate the results are outside of a linear range.

However, if the systems were capable of supporting a records audit, data manipulations by the system were always found to be correct, and quality control flags set by the system were in agreement with original results in almost every case.

Conclusions

The results of the Survey of Laboratory Automated Data Management Practices data evaluation revealed that in the majority of areas surveyed the respondent laboratories were in most cases in compliance with many of the already established GLPs and standards for automation. However, four areas were identified as having substantial deficiencies with respect to meeting the current standards for EPA laboratories. Those areas are independence of quality assurance personnel, system security, system documentation, and practices of data editing.

According to the GLPs, a quality assurance unit or group must be independent of day-to-day laboratory operations to provide an unbiased review of the quality of work conducted. In many of the laboratories, the quality assurance function was not independent, in that the individual responsible for quality assurance typically reported to the laboratory chief; in some of the laboratories, that individual was the day-to-day laboratory manager.

As mentioned previously, of all the areas evaluated by the survey, system security possessed the most variability and lack of continuity between respondent laboratories. More than 90 percent of the systems had not been subjected to a formal risk assessment.

None of the laboratories surveyed had a full complement of the system documentation. Interestingly enough, roughly 65 percent of the laboratories surveyed felt their available documentation represented a conscientious effort to document the .system software. Clearly, in the area of system documentation there exists a need to educate laboratory personnel in the steps necessary to achieve compliance with system documentation standards.

Procedures used for data changes or editing were often in conflict with GLP requirements, which state that data editing must be clearly documented to include when changes were made, who made the changes, and why. More than four out of five respondents indicated that there was no written documentation requirement for who requested or who made the change to data in the system, and less than half reported that their systems kept a log of the data change information.

The Agency should assume responsibility for establishing standards for safeguarding security of computer-resident laboratory data. Sound data offer a fundamental resource for EPA's mission to protect the public health and the environment. The need for standards and guidance is recognized by the laboratories; at every laboratory to which a subsequent on-site visit was made to confirm the findings of this investigation, laboratory staff asked EPA for assistance.

GLOSSARY

Application controls - one of the two sets or types of controls recognized by the auditing discipline. They are specific for each application and include items such as data entry verification procedures (for instance, re-keying all input); data base recovery and roll back procedures that permit the data base administrator to recreate any desired state of the data base; audit trails that not only assist the data base administrator in recreating any desired state of the data base, but also provide documentary evidence of a chain of custody for data; and use of automated reconciliation transactions that verify the final data base results against the results as reconstructed through the audit trail.

Application software - a program developed, adapted, or tailored to the specific user requirements for the purpose of data collection, data manipulation, data output, or data archiving [Drug Information Association].

Audit trail - records of transactions that collectively provide documentary evidence of processing, used to trace from original transactions forward to related records and reports or backwards from records and reports to source transactions. This series of records documents the origination and flow of transactions processed through a system [Datapro]. Also, a chronological record of system activities that is sufficient to enable the reconstruction, reviewing, and examination of the sequence of environments and activities surrounding or leading to an operation, a procedure, or an event in a transaction from its inception to final results [NCSC-TG-004].

Auditing - (1) the process of establishing that prescribed procedures and protocols have been followed; (2) a technique applied during or at the end of a process to assess the acceptability of the product. [Drug Information Association]; (3) a function used by management to assess the adequacy of control [Perry]. That is, auditing is the set of processes that evaluate how well controls ensure data integrity. As a financial example, auditing would include those activities that review whether deposits have been attributed to the proper accounts; for example, providing an individual with a hard-copy record of the transaction at the time of deposit and sending the individual a monthly statement that lists all transactions.

Automated laboratory data processing - calculation, manipulation, and reporting of analytical results using computer-resident data, in either a LIMS or a personal computer.

Availability - see "data availability."

Back-up - provisions made for the recovery of data files or software, for restart of processing, or for use of alternative computer equipment after a system failure or disaster [Drug Information Association].

Change control - ongoing evaluation of system operations and changes during the production use of a system, to determine when and if repetition of a validation process or a specific portion of it is necessary. This includes both the ongoing, documented evaluation, plus any validation testing necessary to maintain a product in a validated state [Drug Information Association].

Checksum - an error-checking method used in data communications in which groups of digits are summed, usually without regard for overflow, and that sum checked against a previously computed sum to verify that no data digits have been changed [Drug Information Association].

Cipher - a method of transforming a text in order to conceal its meaning.

Confidentiality - see "data confidentiality."

Control - "that which prevents, detects, corrects, or reduces a risk" [Perry, p.45], and thus reasonably ensures that data are complete, accurate, and reliable. For instance, any system that verifies the sample number against sample identifier information would be a control against inadvertently assigning results to the wrong sample.

Computer system - a group of hardware components assembled to perform in conjunction with a set of software programs that are collectively designed to perform a specific function or group of functions [Drug Information Association].

Data - a representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by human or automatic means [ISO, as reported by Drug Information Association].

Data availability - the state when data are in the place needed by the user, at the time the user needs them, and in the form needed by the user [NCSC-TG-004-88]' the state where information or services that must be accessible on a timely basis to meet mission requirements or to avoid other types of losses [OMB]. Data stored electronically require a system to be available in order to have access to the data. Data availability can be impacted by several factors, including system "down time," data encryption, password protection, and system function access restriction.

Data Base Management System (DBMS) - software that allows one or many persons to create a data base, modify data in the data base, or use data in the data base (e.g., reports).

Data base - a collection of data having a structured format.

Data confidentiality - the ability to protect the privacy of data; protecting data from unauthorized disclosure [OMB].

Data element (field) - contains a value with a fixed size and data type (see below). A list of data elements defines a data base.

Data integrity - ensuring the prevention of information corruption [modified from EPA Information Security Manual]; ensuring the prevention of unauthorized modification [modified from OMB]; ensuring that data are complete, consistent, and without errors.

Data record - consists of a list of values possessing fixed sizes and data types for each data element in a particular data base.

Data types - alphanumeric (letters, digits, and special characters), numeric (digits only), boolean (true or false), and specialized data types such as date.

Electronic data integrity - data integrity protected by a computer system; automated data integrity refers to the goal of complete and incorruptible computer-resident data.

Encryption - the translation of one character string into another by means of a cipher, translation table, or algorithm, in order to render the information contained therein meaningless to anyone who does not possess the decoding mechanism [Datapro].

Error - accidental mistake caused by human action or computer failure.

Fraud - deliberate human action to cause an inaccuracy.

General controls - one of the two sets or types of controls recognized by the auditing discipline. These operate across all applications. These would include developing and staffing a quality assurance program that works independently of other staff; developing and enforcing documentation standards; developing standards for data transfer and manipulation, such as prohibiting the same individual from both performing and approving sample testing; training individuals to perform data transfers; and developing hardware controls, such as writing different backup cycles to different disk packs and developing and enforcing labelling conventions for all cabling.

Integrity - see "data integrity."

Journaling - recording all significant access or file activity events in their entirety. Using a journal plus earlier copies of a file, it would be possible to reconstruct the file at any point and identify the ways it has changed over a specified period of time [Datapro].

Laboratory Information Management System (LIMS) - automation of laboratory processes under a single unified system. Data collection, data analysis, and data reporting are a few examples of laboratory processes that can be automated.

Password - a unique word or string of characters used to authenticate an identity. A program, computer operator,or user may be required to submit a password to meet security requirements before gaining access to data. The password is confidential, as opposed to the user identification [Datapro].

Quality assurance - (1) a process for building quality into a system; (2) the process of ensuring that the automated data system meets the user requirements for the system and maintains data integrity; (3) a planned and systematic pattern of all actions necessary to provide adequate confidence that the item or product conforms to established technical requirements [ANSI/IEEE Std 730-1981, as reported by Drug Information Association].

Raw data - ". . . any laboratory worksheets, records, memoranda, notes, or exact copies thereof, that are the result of original observations and activities of a study and are necessary for the reconstruction and evaluation of that study. . . "Raw data" may include photographs, microfilm or microfiche copies, computer printouts, magnetic media, . . . and recorded data from automated instruments." [40 CFR 792.3] Raw data are the **first** or **primary** recordings of observations or results. Transcribed data (e.g., manually keyed computer-resident data taken from data sheets or notebooks) are **not** raw data.

Risk - "the probable result of the occurrence of an adverse event..." [Perry, p.45]. An "adverse event" could be either accidental (error) or deliberate (fraud). An example of an adverse event would be the inaccurate assignment of an accessionary number to a test sample. Risk, then, would be the likelihood that the results of an analysis would be attributed to the wrong sample.

Risk analysis - a means of measuring and assessing the relative vulnerabilities and threats to a collection of sensitive data and the people, systems, and installations involved in storing and processing those data. Its purpose is to determine how security measures can be effectively applied to minimize potential loss. Risk analyses may vary from an informal, quantitative review of a microcomputer installation to a formal, fully quantified review of a major computer center [EPA IRM Policy Manual].

Security - the protection of computer hardware and software from accidental or malicious access, use, modification, destruction, or disclosure. Security also pertains to personnel, data, communications, and the physical protection of computer installations [Drug Information Association].

System - (1) a collection of people, machines, and methods organized to accomplish a set of specific functions; (2) an integrated whole that is composed of diverse, interacting, specialized structures and subfunctions; (3) a group of subsystems united by some interaction or interdependence, performing many duties but functioning as a single unit [ANSI N45.2.10, 1973, as reported by Drug Information Association].

System Development Life Cycle (SDLC) - a series of distinct phases through which development projects progress. An approach to computer system development that begins with an evaluation of the user needs and identification of the user requirements and continues through system design, module design, programming and testing, system integration and testing, validation, and operation and maintenance, ending only when use of the system is discontinued [modified from Drug Information Association].

Transaction log - also **Keystroke, capture, report, and replay** - the technique of recording and storing keystrokes as entered by the user for subsequent replay to enable the original sequence to be reproduced exactly [Drug Information Association].

Valid - having legal strength or force, executed with proper formalities, incapable of being rightfully overthrown or set aside [Black's Law Dictionary].

Validity - legal sufficiency, in contradistinction to mere regularity (being steady or uniform in course, practice, or occurrence) [Black's Law Dictionary].

References

Black, Henry C. (1968), *Black's Law Dictionary,* Revised Fourth Edition (West Publishing Co., St. Paul, Minnesota).

Datapro Research (1989), *Datapro Reports on Information Security* (McGraw-Hill, Inc., Delran, New Jersey).

Drug Information Association (1988), *Computerized Data Systems for Nonclinical Safety Assessment: Current Concepts and Quality Assurance* (Drug Information Association, Maple Glen, Pennsylvania).

National Bureau of Standards (1976), *Glossary for Computer Systems Security* (U.S. Department of Commerce, FIPS PUB 39).

National Computer Security Center (1988), *Glossary of Computer Security* (U.S. Department of Defense, NCSC-TG-004-88, Version 1).

Office of Information Resources Management (1987), *EPA Information Resources Management Policy Manual,* Chapter 8 (U.S. Environmental Protection Agency, Washington, D.C.).

Office of Information Resources Management (1989a), *Survey of Laboratory Automated Data Management Practices* (U.S. Environmental Protection Agency, Washington, D.C., April 21, 1989).

Office of Information Resources Management (1989b), *EPA System Design and Development Guidance,* Vols. A, B, and C (U.S. Environmental Protection Agency, Washington, D.C., June 1989).

Office of Information Resources Management (1989c), *EPA Information Security Manual* (U.S. Environmental Protection Agency, Washington, D.C., December 15, 1989).

Office of Management and Budget (1988), *Guidance for Preparation and Submission of Security Plans for Federal Computer Systems Containing Sensitive Information*, OMB Bulletin No. 88-16 (Office of Management and Budget, Washington, D.C., July 6, 1988).

Perry, William E. (1983), *Ensuring Data Base Integrity* (John Wiley and Sons, New York).

U.S. Environmental Protection Agency (1989a), *Federal Register.* Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA); Good Laboratory Practice Standards; Final Rule. 40 CFR Part 160. Vol. 54, No. 158, 34052-74.

U.S. Environmental Protection Agency (1989b), *Federal Register.* Toxic Substance Control Act (TSCA); Good Laboratory Practice Standards; Final Rule. 40 CFR Part 792. Vol. 54, No. 158, August 17, 1989, 34034-50.

APPENDIX A

Detailed Findings from Laboratory Survey

SURVEY OF LABORATORY AUTOMATED DATA MANAGEMENT PRACTICES

Findings from the survey are presented in three columns: the percent of respondents that indicated "yes" to specific questions, the percent that indicated "no," and the overall response rate to each specific question (coded "RESP." on the following pages).

Submitted To:

Office of Information Resources Management
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

April 21, 1989

Submitted By:

Research and Evaluation Associates, Inc.

Contract No. 68-02-4546
Task 10

1030 15th Street, N.W., Suite 750
Washington, D.C. 20005
(202) 842-2200

100 Europa Drive, Suite 590
Chapel Hill, N.C. 27514
(919) 968-4961

NOTICE

This document is a survey distributed by the U.S. Environmental Protection Agency (EPA) to collect information about automated laboratory practices. It does not reflect EPA policy or operational standards and should not be quoted or cited as such. This report has been prepared for EPA's Office of Information Resources Management (OIRM) by Research and Evaluation Associates, Inc.

All responses and/or comments on the survey should be directed by July 21, 1989 to:

Mr. Rick Johnson
U.S. Environmental Protection Agency
Office of Information Resources Management (MD-30)
Research Triangle Park, NC 27711
FTS 629-1132
Commercial (919) 541-1132

EXECUTIVE SUMMARY

The mission of the U.S. Environmental Protection Agency (EPA) is to protect public health and the environment from unreasonable risk. The scientific and technical measurements fundamental to this mission must be accurate and of sufficient integrity to withstand legal scrutiny. As a result, Good Laboratory Practices (GLPs) have been developed for laboratories to follow which assure that measurements will be reliable. EPA has well developed procedures and practices for inspection of laboratories to evaluate compliance with the Agency's GLPs.

The computer has become an accepted resource in the laboratory. Both industry and government laboratories are moving towards more reliance upon computer technology to manage laboratory operations and to interface with laboratory equipment and generate scientific/technical reports. The Agency is rapidly moving to this technology.

Reliance on computer technology in the laboratory has generated a need for a new area of expertise and support to assure that computer resident data are accurate and defensible. The Agency recognizes this need and has initiated a program designed to insure that computer resident data are reliable and defensible. The first phase of this program involves the use of this document to survey laboratories which rely on automated technology.

This document is developed, to a significant degree, from existing documents and standards for automated operations and also from published GLP requirements. The Computer Security Act, the Agency's Data Transmission Standards, the Agency's guidelines for System Design and Development, and the GLPs published by EPA's Office of Toxic Substances and its Office of Pesticide Program have been extensively relied upon in formulating the checklist in this survey. As such, it is an amalgamation of these requirements. Limited interpretations of these published requirements were necessary to develop questions in this checklist. The Agency will evaluate their interpretation of these documents when the questions are used in its upcoming site visits to collect information about data management practices in automated laboratories.

Subsequently, the Agency plans to develop and promulgate standards for automated laboratories. The Agency also intends to examine the development of evaluation criteria to use in auditing automated laboratories; the training and certification of laboratory auditors; and development and implementation of a program for assuring laboratory compliance.

TABLE OF CONTENTS

I

SURVEY PLAN

Introduction

Assuring the integrity of computer resident data demands an integrated laboratory program combining three key elements: staff, documentation, and operations. This document provides a checklist to examine data management practices as part of a survey of laboratories providing data to EPA. It is designed to determine the integrity of computer resident data and relies upon the following assumptions.

The staff must have demonstrated experience and/or training in data management operations. The staff is responsible for the design and implementation of procedures to ensure that data integrity is protected. Ideally, one individual in the laboratory will be designated as the "responsible person". This individual is responsible for ensuring that the data output from the system is an accurate reflection of the data input. This individual is responsible for ensuring that the laboratory is appropriately staffed, has properly documented procedures, and that the procedures are implemented throughout the operation.

Documentation of procedures is required. It assures data integrity by providing a consistent reference to be used by the staff. Documentation of procedures should be based upon good laboratory practices (GLPs) and thus relies on established scientific protocol for ensuring that data can be traced from its source to its final output.

Laboratory operations that ensure data integrity are the sum of well-planned procedures executed by experienced personnel. Even the most well-written documentation will not ensure data integrity unless the laboratory personnel operate within stated procedures. The actual laboratory operations should demonstrate that data can be traced from its entry into the data base, through all data manipulations, and to its final output.

This checklist was devised after review of documents indicated in the bibliography provided as part of the Statement of Work and other documents provided by the staff of the Office of Information Resources Management. The concepts are based upon good laboratory practices as defined by the Toxic Substances Control Act (TSCA), Federal Insecticide, Fungicide and Rodenticide Act (FIFRA), and the Food and Drug Administration (FDA). Additional information was found through review of the Computer Security Act of 1987, the U. S. Environmental Protection Agency (EPA) System Design and Development Guidance, EPA's Data Standards for the Electronic Transmission of Laboratory Measurement Results, and an intensive literature search.

This checklist was developed under Task 10 of Contract Number 68-02-4546 and was designed to be used by a non-computer literate survey respondent. The checklist guides the respondent through four key phases including organization, documentation, operations, and traceability. The following areas will be assessed, in each of these phases:

1) Organization

The objective of this phase is to evaluate the adequacy of the organization of the laboratory data management system(s) and its staff.

- o General description of the data system
- o Data processing personnel
- o Quality assurance personnel
- o Training
- o Security

2) Documentation

The objective of this phase is to establish the adequacy of the documentation for completeness.

- o Written procedures for design, testing, implementation, use, and maintenance of the data processing system

3) Operations

The objective of this phase is to determine that actual laboratory operations directly related to the data system(s) comply with GLPs.

- o Audit trails for data entry and changes
- o Procedures for data input and output
- o System operations including backups and maintenance

4) Traceability

The objective of this final phase assess data integrity.

- o Adequacy of system to aid in performance of traceability study
- o Records audit

II

GUIDE TO USING THE CHECKLIST

This checklist is to be used as a guideline to perform a survey of the data system(s) used in automated laboratories. It is recommended that the survey be conducted by the quality assurance manager or someone not directly responsible for laboratory and/or systems operations (hereafter referred to respondent). The goal is to get an accurate picture of the actual laboratory management practices that is unbiased by assumptions about procedures. The survey is conducted in four phases.

1) Organization

The first is to be used during by key laboratory personnel (designated responsible person and the quality assurance manager). This first phase should provide an overview of how the laboratory is to operate according to its management. This will include a general description of the data processing system and descriptions of the responsibilities of key personnel, training procedures, and security procedures.

2) Documentation

The next major section of the checklist is used as a benchmark of the available written documentation in the laboratory concerning data management operations. This review will enable an evaluation of the completeness and organization of the written documentation. It will also serve as a reference source to use in evaluating the actual laboratory operations.

3) Operations

In this next phase, the checklist is used during an extensive examination of the laboratory operations. Through interviews with the staff and by direct observation of procedures and records, the respondent will evaluate whether the actual laboratory operations reflect those given by management and written procedures.

4) Traceability

The final phase of the survey evaluates data elements from input to output to support a determination of the data integrity. This can be viewed as an internal audit of data. The respondent will track selected data elements through the system and verify the accuracy of data input, the ease of data retrieval and cross-referencing, the adequacy, reasonableness and accuracy of flagging criteria, the accuracy of data conversions and/or manipulations, and the accuracy of data outputs.

III

ORGANIZATION

System Overview

Purpose: To evaluate the adequacy of the automated data processing system(s) used in the laboratory.

Reference: Computer Security Act, (Public Law 100-235, 1988).

Methodology: It is particularly important during this initial phase to determine how many systems are involved in data processing. This will determine how the rest of the checklist will be used throughout the survey. If it is determined that one single system exists then the remaining portion of the checklist is to be administered as it has been provided. If it is determined that more than one system exists, then the checklist will have to be completed for each major operating system in the laboratory. A few examples are provided to aid in assigning the number of systems to be evaluated.

A single system would probably be assigned to a laboratory with one central data system for data entry, tracking, capture from analytical instruments, and data output. A single system might be assigned, for example, in a laboratory using a Laboratory Information Management System (LIMS) on a Perkin-Elmer computer that captures both manually entered incoming sample information and data transferred directly from a gas chromatograph/mass spectrometer (GC/MS).

Two systems might be designated in a laboratory where two management groups are responsible for separate aspects of sample processing being performed on two systems. For example, a central system under the management of the systems group might be used to create sample records used to capture receiving information, reduced test data, and for final reporting. The laboratory manager might be responsible for another system, such as a Hewlett-Packard data server or a series of personal computers, performing data reductions. Each of these systems would require appropriate staff, documentation, and operational guidelines, and would be evaluated independently.

In another laboratory, two operational systems might be resident upon a single central computer. For example, the systems group might be responsible for the data capture software, maintenance of hardware, and system back-ups. The laboratory manager might use the same system, with independently designed and operated software, to reduce analytical data. A system would be designated for each of the two separate data processing functions using software managed by the two distinct groups.

Using the first section of this checklist, the number of computer systems (or the level of aggregation) are determined. Each unique system is defined as a single identifiable system or a group of similar systems having sufficiently similar characteristics/functions to be managed as a single system. For the purposes of this checklist, an analytical instrument, such as an automated gas chromatograph or GC/MS, is not to be considered a separate system.

System Identification

	YES	NO
1) Is there a flow diagram available in the laboratory that gives an overview of the system(s)? If yes, please attach.	40.9	50.1
2) If no flow diagram is available, attach a sketch of the system.		
3) Review the flow diagram provided or constructed. Is the system described either		
a) one identifiable system performing all data processing/sample handling?	62.5	37.5
b) or, one group of similar systems having sufficiently similar functions as to be managed as a single system?	26.7	73.3
c) or, on one identifiable system performing one of two or more clearly distinct and separately managed data handling functions?	15.4	34.6
4) Is there more than one identifiable system, group of systems, or clearly distinct data handling function within a single computer? If yes, briefly describe the separate functions on each system:________________	33.3	66.7

____

____

____

____

____

____

____

____

INSTRUCTIONS FOR PROCEEDING: If the answer to questions 3a, b, or c is yes, then proceed with the checklist as provided in pages 7 to 48. If the answer to question 4) is yes, then for each separately identified system, ask that a copy be made of this checklist for each system and proceed to answer the same questions for each system.

System Environment

	YES	NO	RE
1) Estimate the average number of system operation hours due to power outages. 1 3 Hours per year	13	AVG	88
2) Is the system in an area that provides easy access by			
a) operators?	100.0	0.0	100
b) maintenance staff?	96.0	4.0	100
3) Is the area clean and uncluttered?	96.0	4.0	100
4) Are there precautions taken to ensure continuous system operation such as			
a) is the system located in a temperature-controlled environment?	96.0	4.0	100
b) are there surge protectors used to connect the equipment to power supplies?	83.3	16.7	96
c) is an uninterruptable power supply (UPS) connected?	24.0	76.0	100
d) is the system exposed to any corrosive atmospheric conditions? (such as the exhaust from an atomic absorption instrument)?	16.0	84.0	100
e) are halon devices available to protect system from fire?	64.0	36.0	100

System Description

1) What kind of system is in use?
(Describe manufacturer and model)

Manufacturer:______________________________
____

Model:____________________________________

		YES	NO
2)	Does the system		
	a) allow input of sample receipt information?	80.0	20.0
	b) track samples?	72.0	28.0
	c) do workload scheduling?	40.0	60.0
	d) capture analytical data?	66.7	33.3
	e) perform data reduction?	62.5	37.5
	f) link quality control samples to to case samples?	62.5	37.5
	g) flag samples that fail quality control criteria?	41.7	58.3
	f) allow quality control monitoring?	58.3	41.7
	g) generate quality control reports?	41.7	58.3
	h) generate quality control charts?	29.2	70.8
	i) allow data base management?	92.0	8.0
	j) generate in-house reports?	84.0	16.0
	k) generate client-deliverable reports?	68.0	32.0
	l) provide customer information for in-house use?	62.5	37.5
	m) allow clients access to information via direct computer links?	20.0	80.0
	n) do client billing?	29.2	70.8
	m) provide for electronic transmission of computer-readable data, either by direct link or via magnetic media?	73.9	26.1
3)	Was the data management system software		
	a) developed in-house?	45.4	54.6
	b) provided by the manufacturer?	59.1	40.9
	c) provided by the manufacturer and modified in-house?	60.0	40.0

Personnel

Purpose: To evaluate the qualifications of key personnel responsible for data management and data quality assurance.

Reference: TSCA GLPs (52 FR 48933, 1987)
FIFRA GLPs (52 FR 48920, 1987)

Methodology: This section of the checklist is to be used by two key individuals: the system administrator and the quality assurance manager. The system administrator may be the manager of the systems group or the laboratory or technical manager whose responsibilities include data system management. In any case, it is essential that the laboratory identify an individual who is being held ultimately responsible for the integrity of the data base.

It is also important that the laboratory has identified a quality assurance manager that acts as an independent agent to monitor data integrity. Ideally it is this individual who will use the survey and it can be used as an internal auditing tool.

To determine the qualifications and sufficiency of personnel, such issues as types of responsibilities, training, experience, and number of staff must be considered. It may not be possible, at this point, to determine if the staff is fully sufficient. This may become clear as interviews with other members of the operations staff are done to determine, for example, how much overtime is worked.

Responsible Person

YES NO

1) Identify the individual responsible for the data management system:
Name:________________________________
Title:________________________________

2) How many years of experience has this individual had in the management of data systems? State number of years of formal training in automated data processing (ADP).

□□ Number of years 7.0 AVG

□□ Number of years formal ADP training 2.9 AVG

3) Is this individual responsible only for data management? 24.0 76.0
If no, describe the individual's additional duties:____________________
____
____
____
____
____
____

4) If the individual is not only responsible for data management, on the average, how many hours per week does this individual spend on data management responsibilities?

□□ Number of hours 13.7 AVG

5) Is this individual responsible for the following functions?

a) day-to-day management of the data processing group? 79.2 20.8

b) ensures that there are sufficient personnel with adequate training and experience to supervise and conduct data processing functions? 80.0 20.0

c) ensures the continued competency of data processing staff by documentation of their training, review of work performance, and verification of required skills? 70.8 29.2

	YES	NO	RESP.
d) ensures tne oata management system operations have procedure manuals and other documentation which are complete, current, available to all staff, and properly executed by staff?	83.3	16.7	96
e) approves by review and signature all significant changes to written procedures?	54.6	45.5	88
f) establishes procedures for acceptance testing for any changes made to the software?	69.6	30.4	9
g) assures that data is accurately recorded in the data base?	63.6	36.4	88
h) ensures that problems potentially affecting data quality/integrity are			
i) noted when they occur	91.7	8.3	96
ii) subject to documented corrective action?	83.3	16.7	96
6) During how many hours or shifts does the data system operate? ________________	19.5	AVG	96
7) Are there data processing personnel working during all operational hours?	21.7	78.3	9
8) Are there supervisory personnel available during all hours or shifts of operation?	37.5	62.5	9
9) How many years of supervisory experience does the supervisor have? □□ Number of years	6.3	AVG	8
10) How many years of experience does the the supervisor have in data processing? □□ Number of years	7.8	AVG	8
11) Are system operators available all hours of operation?	29.2	70.8	9
12) Are personnel available to answer user questions and resolve problems during all hours of operation?	26.1	73.9	9

Quality Assurance Personnel

YES NO

1) Identify the individual responsible for quality assurance for the laboratory:
Name:______________________________________
Title:______________________________________

2) How many years has this individual been working in the area of quality control and/or quality assurance?

□□ Number of years 5.8 AVG

3) Is this individual involved in quality assurance for the entire laboratory? 94.7 5.3

4) Do these responsibilities specifically include data processing operations? 57.9 42.1

5) Does this individual use the data management system
- a) as an integral part of his/her job? 36.8 63.2
- b) on a daily basis? 38.9 61.1
- c) for report generation? 64.7 35.3
- d) monthly 31.3 68.8
- e) only occasionally? 20.0 80.0

6) Who does this individual report to in the laboratory?
Name:______________________________________
Title:______________________________________

7) Does the individual who supervises the quality assurance group have any direct responsibility for the day-to-day laboratory operation for the EPA work? 36.8 63.2

8) Does the quality assurance group exist as a separate and identifiable entity acting outside the normal laboratory operation for the EPA work? 63.2 36.8

9) Is the quality assurance supervisor responsible for the following:
- a) maintenance of a master copy of of the procedures used by the data processing group? 52.6 47.4
- b) performance of periodic inspections of the laboratory including the data processing operation? 70.6 29.4

	YES	NO	RESP.
c) submission of periodic quality control reports to the Responsible Person identified above noting any problems identified with the data processing and stating the corrective actions taken?	79.0	21.0	76
d) ensures that reported results accurately reflect raw data?	84.2	15.8	76
e) keeps records of inspections and audits?	94.7	5.3	76

Staff Training and Experience

		YES	NO
1)	Are there written job descriptions available for all data management staff?	95.2	4.8
2)	Where are these kept?________________		
3)	When a new staff member begins work in the data processing group does he/she:		
	a) read the written procedures?	85.0	15.0
	b) receive instruction from supervisor concerning job assignments?	100.0	0
	c) receive instruction from co-worker concerning job assignments?	95.0	5.0
	d) perform new assignments in conjunction with trained personnel?	95.0	5.0
	e) perform new assignments under close supervision?	94.7	5.3
	f) perform any tests to demonstrate proficiency with new assignments?	50.0	50.0
	g) have their work reviewed by co-worker?	85.0	15.0
	h) have their work reviewed by supervisor?	100.0	0
4)	Is training of new personnel documented? If yes, where?______________________ ________________________________	23.8	76.2
5)	Are the qualifications of staff documented? If yes, where?______________________ ________________________________	81.0	19.0
6)	Do primary ADP staff qualifications include formal ADP training?	55.0	45.0
7)	Are there opportunities for continuing education or training of personnel?	95.2	4.8
8)	Is this ongoing training documented? If yes, where?______________________ ________________________________	57.1	42.9

Security

Purpose: To evaluate the adequacy of security for computer resident data.

Reference: Computer Security Act (Public Law 100-235, 1988)
TSCA GLPs (52 FR 48933, 1987)
FIFRA GLPs (52 FR 48920, 1987)

Methodology: This section of the checklist is designed for discussion with the system administrator during the initial interview. First, the security needs of the system are addressed determining the requirements for confidentiality, integrity, and availability. The adequacy of data security is assessed and the ways in which security has been managed by the laboratory is evaluated. Among the areas addressed is monitoring the individuals responsible for data entry and any subsequent data changes as required by GLPs.

Security Needs and Risk Assessment

YES NO

1) Does the data system contain information which requires protection from unauthorized disclosure (i.e., is data confidential?) 92.0 8.0

2) Does the system contain information which must be protected from unauthorized modification (i.e., must data integrity be protected)? 79.2 20.

3) Does the system perform time-critical functions that require that the data's availability be protected such as
   a) sample tracking/scheduling critical prompt sample handling? 72.0 28.
   b) monitoring of quality control data which must be reviewed before data can be released for reporting? 43.5 56.
   c) report generation essential for timely submission of data? 84.0 16.
   d) other (describe):______________
   ____
   ____

4) Of the categories listed above (confidentiality, integrity, and availability) indicate by a check mark if the need is primary (P), secondary (S), or of minimal (M) concern:
   a) confidentiality:
      (P) _____ 62.5
      (S) _____ 29.2
      (M) _____ 8.3
   b) integrity:
      (P) _____ 95.8
      (S) _____ 4.2
      (M) _____ --
   c) availability:
      (P) _____ 75.0
      (S) _____ 25.0
      (M) _____ --

(NOTE: level of security concerns may apply to more than one category)

	YES	NO	RESP.
5) Were the risks and associated protection requirements determined by			
a) formal risk analysis?	9.1	90.0	88
b) other means?	45.4	54.6	88
If by other means, please describe:________________			
6) Were specific standards or other guidance used in the design or implementation of security measures?	8.0	92.0	100
If yes, what reference?________________			
7) Is the management of the laboratory involved in			
a) setting security policy?	68.0	32.0	100
b) assignment of security responsibility?	68.0	32.0	100
c) risk/sensitivity assessment?	64.0	36.0	100
d) personnel selection and screening?	84.0	16.0	100
8) Are there development controls such as			
a) procedures for limited use of a system during development?	50.0	50.0	88
b) design review and testing?	59.1	40.9	88
c) certification of software?	45.4	54.6	88
d) protection against running development software against the active data base?	52.2	47.8	92
9) Are there day-to-day procedures to protect operational application systems such as			
a) production controls?	22.7	77.3	88
b) contingency planning?	34.8	65.2	92
c) variance detection (auditing)?	30.4	69.6	92
10) Is there training in security procedures?	38.1	61.9	84
If yes,			
a) do new personnel receive training when they begin work?	50.0	50.0	72
b) a annual refresher courses given to existing personnel?	0	100	72
11) Is physical access to the system limited			
a) locating the system within a secured facility?	92.0	8.0	100

	YES	NO	RESI
b) by securing the facility with a building guard?	44.0	56.0	100
c) locating system within a secured room?	54.2	45.8	96
d) by using cipher locks			
i) for the laboratory?	21.7	78.3	92
ii) for computer room?	26.1	73.9	92
e) by using are card keys			
i) for the laboratory?	22.7	77.3	88
ii) for the computer room?	23.8	76.2	84
f) by other means (describe):______			
12) Are personal computers kept in offices or areas which are locked			
a) during the day?	12.5	87.5	96
b) at night?	54.2	45.8	96

System Access Security

		YES	NO	RESP
1)	Does the system require personalized log-ons for each user?	56.0	44.0	100
2)	Does each user have a password?	56.0	44.0	100
3)	Are there any group user-identifications or passwords used by members of a functional group?	48.0	52.0	100
4)	How often are passwords changed? ____________	67.5 days average		32
5)	Are there established password standards?	36.0	64.0	100
6)	Is access to parts of the system restricted to certain users? *If yes, describe the procedure for* authorization of users:____________	66.7	33.3	96
7)	Are there procedures for protecting the system from the introduction of computer viruses such as			
	a) controls on the number of personnel allowed to introduce external software?	58.3	41.7	96
	b) tracking any external software introduced?	43.5	56.5	92
	c) other (descirbe): ____________			
8)	Does the data management system track changes to the data?	52.0	48.0	100
9)	Does the system automatically flag data as having been edited?	48.0	52.0	100
9)	Is there a record maintained of the unaltered data? If yes, how long is this maintained: ____________	45.8	54.2	96

IV

DOCUMENTATION

Purpose: To determine the adequacy of written procedures for the data management operation.

Reference: TSCA GLPs (52 FR 48933, 1987)
FIFRA GLPs (52 FR 48920, 1987)
EPA System Design and Development Guidance (OIRM, 1987)

Methodology: This section of the checklist serves as an inventory of the types of documentation that are available in a laboratory.

It is assumed that the respondent is familiar with the standard operating procedures of the laboratory. This portion of the survey is to collect information about the kinds of documents that laboratories typically use. It can also serve as a review of written procedures that can later be cross-checked against actual laboratory practices observed during the remainder of the survey.

Not all of the documents listed in the following sections may actually be available in all laboratories. Operations, maintenance, and user's guides, should be present in all laboratories. If the laboratory has designed its own software, it is more likely to have requirements and design documents.

Documentation: The following paragraphs provide a brief description of the types of documents that might be reviewed:

System Implementation Plan: Identifies a project plan for implementation of software. This includes the events, actions, milestones, resources, schedules, and workplans identified for successful completion of the project.

System Detailed Requirements Document: The plan outlines the requirements that the system needs to fulfill. It describes the major system functions, the requirements for security, quality control, testing, verification, and resources expected over the projected lifecycle of the system. It is during this phase of system development that a "data dictionary" should have been prepared. This is a listing of the data elements and field names providing specifications as to the type (numerical or character) and size of each defined data field.

Software Management Plan: Identifies the organizational structure of the project team and define review responsibilities. Describes risk management, software quality assurance, and development procedures that ensure that systems under development are a separate entity from any operational systems.

Software Test and Acceptance Plan: This plan outlines the procedures for testing the system including the tests that were used, who ran the tests, and forms that were used to document that the tests were completed.

Software Preliminary Design Document: This plan should provide details concerning the design of required system functionalities including inputs, outputs, timing, sequencing and error handling for the system.

Software Detailed Design Document: This plan should provide the exact functional details of the designed system including inputs, outputs, data processing/reduction formulas, conversions, test or quality control test structures, interfaces, and error handling.

Software Maintenance Document: This document outlines procedures for maintaining the integrity of the data base. It should include change request procedures, test procedures for problem resolution, procedures for implementation of changes, configuration management, back-up and archival procedures, and methods for reporting operational status and problems to management.

Software User's Guide: This document should be a step-by-step guide to the user for all data processing functions including data entry and updating, data processing, report generation, problem notification, and training.

System Integration Test Reports: This document provides a reference of system problems, corrective actions taken, and a listing of outstanding problems including evaluation and recommended disposition.

System Documents

1) Which of the following documents are available in the laboratory?

	YES	NO	R
a) System Implementation Plan?	28.0	72.0	10
b) System Detailed Requirements Document?	16.7	83.3	9
c) Software Management Plan?	8.3	91.7	9
d) Software Test and Acceptance Plan?	16.7	83.3	9
e) Software Preliminary Design Document?	12.5	87.5	9
f) Software Detailed Design Document?	12.5	87.5	9
g) Software Maintenance Document?	24.0	76.0	10
h) Software Operations Document?	56.0	44.0	10
i) Software User's Guide?	79.2	20.8	9
j) System Integration Test Reports?	12.5	87.5	9

2) Does the Implementation Plan contain the following elements?

a) purpose? ___ ___
b) references? ___ ___
c) table of contents? ___ ___
d) strategy for acquiring information? ___ ___
e) consideration of systems that may have to be integrated into planned system? ___ ___
f) policy for access to the system? ___ ___
g) assessment of existing hardware/ software? ___ ___
h) target workplans/schedules? ___ ___
i) resource requirements including possible contractors or in-house staff? ___ ___

3) Does the System Detailed Design Document include the following elements?

a) purpose? ___ ___
b) references? ___ ___
c) table of contents? ___ ___
d) system definitions including
i) purpose?
ii) operation concept?
iii) system size and timing requirements?
iv) definitions of any sub-systems needed? ___ ___
e) physical requirements for cooling, electricity, security of facility? ___ ___
f) back-up and disaster recovery? ___ ___
g) quality requirements for
i) reliability (up-time)? ___ ___
ii) maintainability? ___ ___
iii) flexibility for expansion? ___ ___
iv) transportability? ___ ___
h) testing methods and responsibility? ___ ___
i) workplan/schedule? ___ ___

YES NO

4) Does the Software Management Plan include the following elements?

a) purpose? ___ ___

b) references? ___ ___

c) table of contents? ___ ___

d) project resources including

i) personnel? ___ ___

ii) outline of staff responsibilities? ___ ___

e) plan for development of software? ___ ___

f) procedures for back-up/recovery during software development? ___ ___

g) procedures for independent validation of software? ___ ___

h) definitions of interfaces? ___ ___

i) quality assurance procedures such as

i) program monitoring? ___ ___

ii) quality reviews? ___ ___

iii) reporting? ___ ___

iv) review of design, coding, tests? ___ ___

j) testing requirements identifying test procedures and reports? ___ ___

k) definitions of software tools including

i) identification of commercial or reusable in-house software? ___ ___

ii) identification of program language? ___ ___

iii) identification of interface or network software requirements? ___ ___

iv) design and coding standards? ___ ___

l) configuration control (i.e., control, release, and storage of master copies)? ___ ___

5) Does the Software Test and Acceptance Plan include the following elements?

a) purpose? ___ ___

b) references? ___ ___

c) table of contents? ___ ___

d) tests done including

i) test requirements? ___ ___

ii) test management? ___ ___

iii) test schedule? ___ ___

iv) test results? ___ ___

e) tests done of sub-systems? ___ ___

f) user acceptance forms? ___ ___

g) outline procedures for formal acceptance including a test report giving complete test history and results? ___ ___

YES NO

6) Does the Software Preliminary Design Document include the following elements?

a) purpose? ___ ___
b) references? ___ ___
c) table of contents? ___ ___
d) flowchart or text describing functional flow? ___ ___
e) design goals for functions including
i) inputs? ___ ___
ii) outputs? ___ ___
iii) initiation, timing and sequencing of events? ___ ___
iv) interrupts? ___ ___
v) algorithms? ___ ___
vi) error handling? ___ ___
f) design goals for data base design including
i) interrelationships? ___ ___
ii) traceability between data bases? ___ ___
iii) data structure (logical design)? ___ ___

7) Does the Detailed Design Document include the following elements?

a) purpose? ___ ___
b) references? ___ ___
c) table of contents? ___ ___
d) identification of interface for each external software unit of instrument? ___ ___
e) for each system function an identification of the final
i) inputs? ___ ___
ii) outputs? ___ ___
iii) initiation, timing and sequencing of events? ___ ___
iv) interrupts? ___ ___
v) algorithms? ___ ___
vi) error handling? ___ ___
f) for each data base, an identification of the final design for
i) interrelationships? ___ ___
ii) traceability between data bases? ___ ___
iii) data structure (logical design)? ___ ___

YES NO

8) Does the Software Maintenance Document include the following elements?

a) purpose? ___ ___

b) references? ___ ___

c) table of contents? ___ ___

d) maintenance procedures for

i) source code standards? ___ ___

ii) requirements for updating documents? ___ ___

iii) coding review by peers or team leader? ___ ___

iv) requirements for commenting on code? ___ ___

e) procedures for change management including instruction for

i) making change requests ___ ___

ii) testing changes? ___ ___

iii) approving changes? ___ ___

iv) implementation of changes? ___ ___

f) procedures for configuration management? ___ ___

g) how to use technical maintenance tools such as compilers, file comparators, traces/dumps, etc.? ___ ___

h) how to use clerical maintenance tools such as on-line editors, updating data dictionary, recording maintenance? ___ ___

i) management procedures for maintenance such as use of problem reports, status reports, scheduling of changes, etc.? ___ ___

9) Does the Software User's Guide include the following elements?

a) purpose? ___ ___

b) references? ___ ___

c) table of contents? ___ ___

d) description of the system? ___ ___

e) identification of system manager or other individual for questions? ___ ___

f) how to access the system? ___ ___

g) how to generate reports including

i) standard reports? ___ ___

ii) ad-hoc report capabilities? ___ ___

iii) specialized capabilities? ___ ___

iv) printer options and selection? ___ ___

h) how to enter and update data? ___ ___

i) listing of error codes? ___ ___

j) availability of user support/training? ___ ___

YES NO

10) Does the System Integration Test Report include the following elements?

a) purpose? ___ ___

b) references? ___ ___

c) table of contents ___ ___

d) summary of testing? ___ ___

e) test results? ___ ___

f) listing of outstanding incidents? ___ ___

g) evaluation and recommendations for disposition of problems unable to be resolved at this time? ___ ___

11) Does the appearance, format, and content of the document(s) reflect a conscientious effort to document the software? ___ ___

V

OPERATIONS

Purpose: To evaluate adherence of laboratory operations to policies and procedures determined during interviews with the management and review of written procedures.

Reference: TSCA GLPs (52 FR 48933, 1987)
FIFRA GLPs (52 FR 48920, 1987)

Methodology: This section of the checklist provides guidance in the evaluation of actual laboratory operations. It is to be used while touring the laboratory operation. The checklists relies upon direct observations, review of records, and staff interviews to examine laboratory practices.

In most laboratories, the most efficient way to conduct this part of the survey is to ask to be move through the laboratory just as a sample would. This will typically start in the receiving area where data is first entered into the data base to describe a sample. Then the sample will move into some phase of preparation, analysis, quality control checking, review, validation, and finally reporting. Again, in most automated laboratories this will allow interface with the data system operations at several different points.

The survey is also specifically interested in the system operation, maintenance, and archival procedures. This will require visiting where the computer is physically located and discussing such items with the system operator on staff.

Throughout this survey, efforts should be made to allow all levels of staff to tell the respondent you what is involved with their work. One of the ways to accomplish this is to ask open-ended questions of personnel. Ask staff members to describe their jobs, ask to see the kinds of data sheets they use to input data or the reports that they routinely generate, or what records they keep. The checklist serves as a tool for the to determine if all areas of concern have been discussed with personnel. If someone does not specifically discuss a task or a record that seems appropriate to the task, ask the staff member about it. Find out if someone else has that responsibility, make a note of this, and later discuss it with that person. If you simply ask a list of "yes and no" questions, the staff will. quickly learn that typically the right response is "yes"!

Data Entry

	YES	NO	RES
1) Does the individual use a personalized log-on to access the system?	64.0	36.0	100
2) Is there a password required to access the system?	60.0	32.0	100
3) Is the individual entering data			
a) from a hard-copy?	91.3	8.7	92
b) by prompting system to access an existing data file?	68.2	31.8	88
c) prompting the system to access data directly from another system or instrument?	52.2	47.8	92
4) Was the individual trained for any data entry or transfer functions?	91.7	8.3	96
5) During training, did the individual			
a) read the written procedures?	87.0	13.0	92
b) receive instruction from supervisor concerning job assignments?	100	0	92
c) receive instruction from co-worker concerning job assignments?	70.8	29.2	96
d) perform new assignments in conjunction with trained personnel?	91.7	8.3	96
e) perform new assignments under close supervision?	79.2	20.8	96
f) perform any tests to demonstrate proficiency with new assignments?	29.2	70.8	96
g) have their work reviewed by co-worker?	70.8	29.2	96
h) have their work reviewed by supervisor?	87.5	12.5	96
6) Is the screen used for data entry			
a) designed to match the forms used for entering data?	50.0	50.0	96
b) assessed to be convenient by the individual responsible for data entry?	95.6	4.4	92
7) Does the data entry staff experience any delays due to the system that hampers their job?	50.0	50.0	88
If yes, does this			
a) delay further sample processing?	50.0	50.0	56
b) cause the individual to work overtime?	35.7	64.3	56
c) other:________________________			

____

____

		YES	NO	RESP
8)	If data is manually entered from a hard-copy is the data validated by			
	a) re-keying by the same person?	12.5	87.5	96
	b) re-keying from another person?	8.3	91.7	96
	c) review by same person?	58.3	41.7	96
	d) review by another person?	87.5	12.5	96
9)	Does the system alert the data entry personnel if an error is made in data entry (i.e., values out of date range or incorrect flags, etc.)?	87.5	12.5	96
10)	Does the system prevent entry of incorrect or out-of-range data?	50.0	50.0	88
11)	Does the system prompt the individual entering data if there are missing fields?	73.9	26.1	92
12)	Are there any default parameters assumed by the system (i.e., assignments of proper dates or the next accessioning number)?	95.8	4.2	96
13)	Does the system carry over any entered data from one screen to the next in order to minimize entry errors?	73.9	26.1	92
14)	If data is entered into the central data base via a computer readable media does this data contain information concerning			
	a) who or what instrument initially collected the data?	64.7	35.3	68
	b) when the data was collected?	76.5	23.5	68
	c) if applicable, under what conditions the data was collected?	47.1	52.9	68
	d) any applicable quality control data?	52.9	47.1	68
	e) pointers to link case data to its associated quality control data?	58.8	41.2	68
	f) quality control flags indicating the level of data acceptability?	64.7	35.3	68
15)	If data is entered by prompting the system to access a previously existing data file is the data validated by			
	a) a comparison of the number/size of of files undergoing transfer?	31.6	68.4	76
	b) a log maintained documenting which files have been transferred?	31.6	68.4	76

	YES	NO	
c) creating a record of the date and the individual responsible for the data transfer?	27.8	72.2	7
d) An audit proving exact data transfer (periodic)?	22.2	77.8	7
16) If data is entered into the central data system via a direct link from an instrument does the data being transferred contain information concerning			
a) who analyzed the data?	47.1	52.9	6
b) when the data was analyzed (including date and time)?	70.6	29.4	6
c) on what instrument the analysis was performed?	70.6	29.4	6
d) instrument conditions?	47.1	52.9	6
e) any applicable quality control data?	55.6	44.4	7
f) pointers to link case data to its associated quality control data?	61.1	38.9	7
g) quality control flags indicating the level of data acceptability?	61.1	38.9	7
16) If data is entered via a direct link from an instrument is the data validated by			
a) periodic voltage and calibration checks?	40.0	60.0	60
b) standardized sample processing and results comparison?	46.7	53.3	60
c) visual comparison of instrument hardcopy output (where available) versus data base contents?	66.7	33.3	60
d) are there data reasonableness checks either built into the instruments or a part of the data capture system?	64.7	35.3	68

Data Changes

		YES	NO	RESP
1)	When data is manually entered into the data base, if changes are required due to clerical errors are they made by			
	a) data entry operator?	91.3	8.7	92
	b) data entry supervisor?	63.6	36.4	88
	c) systems group?	54.6	45.4	88
2)	Does the data processing staff experience any delays due to the system that impede their work?	28.6	71.4	84
	If yes, does this			
	a) delay further sample processing?	60.0	40.0	40
	b) cause the individual to work overtime?	44.4	55.6	36
	c) other:________________________ ________________________ ________________________			
3)	Was the individual trained for making data changes?	91.7	8.3	96
4)	During training, did the individual			
	a) read the written procedures?	81.8	18.2	88
	b) receive instruction from supervisor concerning job assignments?	95.4	4.6	88
	c) receive instruction from co-worker concerning job assignments?	76.2	23.8	84
	d) perform new assignments in conjunction with trained personnel?	81.8	18.2	88
	e) perform new assignments under close supervision?	77.3	22.7	88
	f) perform any tests to demonstrate proficiency with new assignments?	27.3	72.7	88
	g) have their work reviewed by co-worker?	71.4	28.6	84
	h) have their work reviewed by supervisor?	85.7	14.3	84
5)	When making changes does the individual log-on the system with a personalized password?	50.0	50.0	96
6)	If changes are to be made only by certain supervisory personnel, does the data entry staff ever log-on using their supervisor's password to make the changes?	15.0	85.0	80

	YES	NO	RE.
7) Are the corrections verified?	69.6	30.4	92
If yes, describe how:			
8) When the changes are made to the system, is there hard-copy documentation of			
a) who made the change?	17.4	82.6	92
b) when the change was made?	17.4	82.6	92
c) who authorized the change?	8.7	91.3	92
9) Once this initial data is checked and corrected for any clerical errors, is the data committed to the data base?	95.4	4.6	88
10) If the data is committed to the data base			
a) can further changes be made to the data?	86.4	13.6	88
b) is there written documentation for			
i) who requested the change?	14.3	85.7	84
ii) who authorized the change?	9.5	90.5	84
iii) who made the change?	14.3	85.7	84
iv) when the change was made?	14.3	85.7	84
11) If a change is made to the committed data base does the system maintain a log of			
a) who made the change?	40.9	59.1	88
b) when the change was made?	47.6	52.4	84
c) a record of both the unchanged and changed data?	22.7	77.3	88
12) If data is entered into the central data base via a data set on a computer readable media			
a) can further changes be made to the data?	12.5	87.5	64
b) who can authorize changes?			
c) is there written documentation for			
i) who requested the change?			
ii) who authorized the change?			
iii) who made the change?			
iv) when the change was made?			
13) If a change is made to the transferred data base does the system maintain a log of			
a) who made the change?	11.19	88.9	72
b) when the change was made?			
c) a record of both the unchanged and changed data?			

		YES	NO	RESP
14)	If changes are made to the data that has been transferred to the central data base are the associated changes also made on the original data source?	44.4	55.6	72
15)	If the answer to 14) is yes, are the changes documented including			
	a) who authorized the change?	12.5	87.5	32
	b) who made the change?	50.0	50.0	32
	c) when the change was made?	50.0	50.0	32
	d) a record of both the original and changed data sets?	28.6	71.4	28
16)	Are changes made to the data in the central data base derived from a direct link with an instrument?	25.0	75.0	80
17)	Are quality control flags set for data using software resident on analytical instruments?			
	If yes,			
	a) are the flags transferred to the central data base?			
	b) can changes to the flags be made on the central data base (i.e., as a result of data review)? If yes, answer question 18)			
18)	When quality control flags originally set by instrument software are changed in the central data base			
	a) who determines that a change needs to be made:____________________ ________________________________			
	b) Is the change authorized? If yes, by whom:________________ ________________________________			
	c) Does the system maintain a record of			
	i) who made the change?			
	ii) when it was made?			
	iii) both the changed and unchanged data flag?			
19)	Are there quality control checks performed by the central data system?	30.0	70.0	80
20)	Do these checks result in setting quality control flags?			
	If yes, once a flag has been set, can it be changed? If yes, answer question 21)			

YES NO

21) When quality control flags are changed in the central data base

a) who determines that a change needs to be made:____________________
____

b) Is the change authorized? ___ ___
If yes, by whom:________________
____

c) Does the system maintain a record of
i) who made the change? ___ ___
ii) when it was made? ___ ___
iii) both the changed and unchanged data flag? ___ ___

Data Reduction, Analysis, and Assessment

		YES	NO	RESP
1)	When data is manually entered are any validation flags set?	41.7	58.3	96
2)	Are there written charts or tables available defining flags?	62.5	37.5	96
3)	Is the chart current?	87.5	12.5	64
4)	Can the data entry person			
	a) create new flags?	0	100	76
	b) request that new flags be added?	57.9	42.1	76
5)	Are the algorithms or formulas used for data manipulations performed by the system available in a written format?	63.6	36.4	88
6)	Were these algorithms or formulas reviewed for accuracy by quality assurance staff?	75.0	25.0	80
7)	Are a minimum of two test data records processed to test each algorithm?	65.0	35.0	80
8)	Are the test results			
	a) documented?	35.0	65.0	80
	b) reviewed by the quality assurance staff?	30.0	70.0	80
9)	Are a minimum of ten data records processed to test each validation algorithm?	36.8	63.2	76
10)	Are the test results			
	a) documented?	29.4	70.6	68
	b) reviewed by the quality assurance staff?	23.5	76.5	68
11)	Are these checks done			
	a) during system development?	76.2	23.8	84
	b) whenever changes are made in the data base?	42.9	57.1	84
	c) periodically by quality assurance staff?	14.3	85.2	84
	d) through the use of internal quality control samples?	33.3	66.7	84

		YES	NO	R
12)	If algorithms or formulas are modified			
	a) is this documented?	63.2	36.8	7
	b) is it possible to determine which data sets were processed with which version of the calculations?	21.0	79.0	7
	c) are old data recalculated with new formulas?	27.8	72.2	7
	d) are changes reflected in the detail design documentation?	10.5	89.5	7
13)	Are computer printouts and reports routinely checked against field and laboratory data before data are released?	85.0	15.0	7

a) by whom?_________________________

____

b) date of last check:______________

c) how many/what percent are checked:

____

Data Outputs

		YES	NO	RES
1)	Are there written procedures in the laboratory for generation of reports, graphs, and charts?	70.8	29.2	96
2)	Do the procedures include			
	a) what information must be input to generate the product?	61.9	38.1	84
	b) where the product will be output?	57.1	42.9	84
	c) where to file or deliver the product?	40.0	60.0	80
3)	Was the individual trained for generating reports?	96.0	4.0	100
4)	During training, did the individual			
	a) read the written procedures?	78.3	21.7	92
	b) receive instruction from supervisor concerning job assignments?	96.0	4.0	100
	c) receive instruction from co-worker concerning job assignments?	79.2	20.8	96
	d) perform new assignments in conjunction with trained personnel?	91.7	8.3	96
	e) perform new assignments under close supervision?	84.0	16.0	100
	f) perform any tests to demonstrate proficiency with new assignments?	29.2	70.8	96
	g) have their work reviewed by co-worker?	75.0	25.0	96
	h) have their work reviewed by supervisor?	91.7	8.3	96
5)	Is the program/screen used for report generation assessed to be convenient by the user?	83.3	16.7	96
6)	Does the data processing staff experience any delays due to the system that hamper their job?	50.0	50.0	88
	If yes, does this			
	a) delay further sample processing?	36.4	63.6	44
	b) cause the individual to work overtime?	36.4	63.6	44
	c) other:________________________ ________________________ ________________________			
7)	When interim reports are generated, is it determined that the data reduction necessary for that report is complete?	75.0	25.0	80

	YES	NO	RESP
8) When final reports are generated			
a) is it determined that the data reduction necessary for that report is complete?	90.5	9.5	84
b) is the data base "locked" so no further changes can be made to the data?	29.2	70.8	96
9) Does the system generate reports on a timely basis?	91.7	8.3	96
If not, how long does the user typically have to wait to a report?________________			
10) Can the user create customized reports?	76.0	24.0	100
11) Can the user request new report formats from the systems group?	91.7	8.3	96
If yes, what is typical response time from the systems group:______________________ ______________________________________			
12) If hard-copies of final reports are archived are they			
a) stored in an off-site location?	43.5	56.5	92
b) stored in a secure area?	88.0	12.0	100
c) stored in a fireproof area?	26.1	73.9	92
d) stored with proper identification to facilitate retrieval?	100	0	96
e) accessible only to designated staff?	83.3	16.7	96
13) Are data reports electronically transmitted?	54.2	45.8	96
If yes, is transmission			
a) by direct computer link?	46.7	53.3	60
b) via magnetic media?	91.7	8.3	48
c) verified?	40.0	60.0	60
Describe how:________________ ___________________________ ___________________________ ___________________________			

Back-ups/Archival

	YES	NO	RESP
1) Are system back-ups performed?	100	0	100
2) If back-ups are performed, how often?			
a) daily?	77.8	22.2	72
b) weekly?	60.0	40.0	60
c) monthly?	76.9	23.1	52
d) other:____________________			
3) Are the back-ups			
a) partial?	76.2	23.8	84
b) total?	91.5	8.7	92
4) Is there a designated individual responsible for system back-ups? If yes, who:____________________	87.5	12.5	96
5) Was the individual trained for making back-ups/archivals?	92.0	8.0	100
6) During training, did the individual			
a) read the written procedures?	81.8	18.2	88
b) receive instruction from supervisor concerning job assignments?	72.7	27.3	88
c) receive instruction from co-worker concerning job assignments?	57.1	42.9	84
d) perform new assignments in conjunction with trained personnel?	72.7	27.3	88
e) perform new assignments under close supervision?	68.2	31.8	88
f) perform any tests to demonstrate proficiency with new assignments?	45.4	54.6	88
g) have his/her work reviewed by co-worker?	50.0	50.0	88
h) have his/her work reviewed by supervisor?	77.3	22.7	88
7) On what media are the back-ups stored			
a) magnetic tapes?	79.2	20.8	96
b) disks?	43.8	56.2	64
c) diskettes?	70.6	29.4	68
d) other:____________________			
8) Are the media storing back-ups properly labelled?	100	0	100

	YES	NO	RESP
9) When the system is backed-up, is this documented?	68.0	32.0	10(
If yes,			
a) documented in a written log?	70.6	29.4	68
b) documented on the system?	58.8	41.2	68
c) other:______________________			
10) Are command files written to drive back-up operations?	76.0	24.0	10(
11) Are back-up media stored for the short-term at the laboratory facility?	91.7	8.3	96
If yes, for how long?__________________			
12) During this short-term storage, are back-up media			
a) stored in a secure area?	91.3	8.7	92
b) stored with proper identification to facilitate retrieval?	95.6	4.4	92
c) accessible to authorized staff only?	69.6	30.4	92
13) For long-term archival, are back-up media			
a) stored in an off-site location?	20.8	79.2	96
b) stored in a secure area?	87.0	13.0	92
c) stored in a fire-proof area?	30.4	69.6	92
d) stored with proper identification to facilitate retrieval?	91.3	8.7	92
e) accessible only to authorized staff?	69.6	30.4	92

System Maintenance

	YES	NO	RESP
1) Is an individual designated as responsible for system maintenance?	90.9	9.1	88
2) Was the individual trained for system maintenance?	70.0	30.0	80
3) During training, did the individual			
a) read the written procedures?	81.2	18.8	64
b) receive instruction from supervisor concerning job assignments?	76.5	23.5	68
c) receive instruction from co-worker concerning job assignments?	58.8	41.2	68
d) perform new assignments in conjunction with trained personnel?	76.5	23.5	68
e) perform new assignments under close supervision?	82.4	17.6	68
f) perform any tests to demonstrate proficiency with new assignments?	52.9	47.1	68
g) have their work reviewed by co-worker?	56.2	43.8	64
h) have their work reviewed by supervisor?	81.2	18.8	64
4) Is there a regularly scheduled preventative maintenance program?	52.6	47.4	76
If yes, what frequency?			
a) weekly?	57.1	42.9	28
b) monthly?	85.7	14.3	28
c) quarterly?	66.7	33.3	24
d) annually?	50.0	50.0	16
e) other:____________________			
5) Is the preventative maintenance documented indicating			
a) what was done?	61.9	38.1	84
b) who did the work?	61.9	38.1	84
c) how long the system was down?	61.9	38.1	84
6) Is non-routine maintenance performed by in-house staff?	59.1	40.9	88
7) Is non-routine maintenance documented indicating			
a) the nature of the problem?	68.2	31.8	88
b) what was done to correct problem?	68.2	31.8	88
c) who performed the work?	72.7	27.3	88
d) how long the system was down?	59.1	40.9	88

Repair Service

		YES	NO	RES
1)	Are there contracted technicians to make repairs?	86.4	13.6	88
2)	Is there a response time designated in the service contract? If yes, what is it?____________________	75.0	25.0	80

7.62 hours Averag

3) What is the typical response time of the repair service:________________________

4.33 hours Averag

4) What provisions are there to continue laboratory operations if the system is down?

____
____
____
____
____

5) If the system is down, what impact does it have on the laboratory operation?____________

____
____
____
____

Recovery From System Failure

		YES	NO	RESP.
1)	If the system fails due to a power failure or glitch does the system			
	a) restart automatically?	42.9	57.1	84
	b) have a manual restart?	100	0	76
	c) other:________________________			
2)	Does the system lose the data being processed?	61.6	38.4	72
3)	Does the system start from where if left off?	27.3	72.7	88
4)	If data is lost, can the system show the loss and identify which data was lost?	21	79	76
5)	Is there a back-up procedure done on a regular basis to minimize data loss?	90.9	9.1	88
6)	Is there a disaster recovery plan for data retrieval?	19.0	81.0	84

VI

TRACEABILITY

Purpose: **To determine if the information in the data system accurately reflects the raw data.**

References:
TSCA GLPs (52 FR 48933, 1987)
FIFRA GLPs (52 FR 48933, 1987)
Data Standards for Electric Transmission of Laboratory Measurement Results (EPA Order 21802, 1987)

Methodology: This can be accomplished by selection of a number of final data sets (resident derived or progeny data) and tracking the data back to parent data. The parent data would likely be the hard-copy forms used for data entry or the hard-copy graphs/printouts from analytical instruments. It should be will verified that any flags are accurately set or that manipulations performed by the system are done correctly.

It is assumed that the analysis performed by the laboratory instruments are correct. Traditional elements of a traceability study such as laboratory notebooks, standard certification and preparation, or instrument maintenance logs will not be reviewed as part of this survey.

To perform the traceability audit, select data that has been completed by the laboratory. Select final sample cases that both that have been reported as passing all quality controls and also some with reported quality control problems. The mission of the survey is to determine if the reported cases reflect the data collected by the laboratory.

Another important aspect of this phase of the survey is to determine the ease of performing a traceability audit. In some laboratories, it may not even be possible to identify all the data elements used in determination of the final results. The lack of ability to perform a traceability audit is a critical finding.

Records Tracking

	YES	NO	RESP
1) Which of the following records are maintained only on the data system?			
a) results of instrument calibrations?	13.6	86.4	88
b) results of instrument blanks?	27.5	72.7	88
c) results of additional quality control samples such as duplicates, spikes, etc.?	18.2	81.8	88
d) laboratory identification of case samples?	31.8	68.2	88
e) flags made associated with problems found during initial samples receipt (such as missing client information, leakage, etc.)?	22.7	77.3	88
d) flags associated with quality control problems?	28.6	71.4	84
e) records of individuals who review data?	23.8	76.2	84
f) any modifications of data flags made by data review staff?	19.0	81.0	84
g) evidence that data review was completed and samples were released for reporting?	23.8	76.2	84
2) Which of the following records are maintained only on hard-copy records?			
a) results of instrument calibrations?	52.4	47.6	84
b) results of instrument blanks?	28.6	71.4	84
c) results of additional quality control samples such as duplicates, spikes, etc.?	33.3	66.7	84
d) laboratory identification of case samples?	19.0	81.0	84
e) flags made associated with problems found during initial samples receipt (such as missing client information, leakage, etc.)?	42.9	57.1	84
d) flags associated with quality control problems?	26.3	73.7	76
e) records of individuals who review data?	63.6	36.4	88
f) any modifications of data flags made by data review staff?	40.0	60.0	80
g) evidence that data review was completed and samples were released for reporting?	50.0	50.0	80
3) If the data system tracks both case samples and their associated quality control samples, is there a pointer used in the system to link the case sample with			
a) standards?	63.2	36.8	76
b) blanks?	63.2	36.8	76
c) instrument calibrations?	61.1	38.9	72
d) instrument conditions?	44.4	55.6	72
e) duplicates?	57.9	42.1	76
f) spikes?	68.4	31.6	76

	YES	NO	RE
g) internal standards in sample?	61.1	38.9	72
h) surrogate standards in sample?	70.6	29.4	68
i) compounds under investigation?	75.0	25.0	64
j) unknown compounds found in sample?	50.0	50.0	64
4) Is it possible using the data system to change any of these key links (i.e., could a case sample be linked to a different quality control set than that with which it was run)?	38.9	61.1	72
If yes, does the system maintain a record			
a) of who made the change?	44.4	55.6	36
b) who authorized the change?	22.2	77.8	36
c) of both the unchanged and changed case/quality control link?	22.2	77.8	36

If the system does not keep a record of any of these items, where is it kept?

5) Are the links established between case samples and their associated quality control samples sufficient to determine without any questions which quality control samples were run with each case sample?	79.0	21.0	76
6) Are any Electric Data Interchange standards in effect now?			
a) ANSI X.12	11.8	88.2	68
b) other	18.8	81.2	64
7) Are standard data formats specified for all databases in the organization?	31.6	68.4	76

If yes, specify:

Records Audit

	YES	NO	RESP
1) Does the system perform any of the following data reduction functions?			
a) linear or quadratic reduction for standard curves?	39.1	60.9	92
b) quantitative analysis for unknowns utilizing formulas derived in a)	34.8	65.2	92
c) flagging of data to indicate			
i) standards outside of quality control acceptance criteria?	38.1	61.9	84
ii) sample results outside linear range?	18.2	81.8	88
iii) sample results below detection limits?	63.6	36.4	88
iv) sample results below reporting limits?	54.6	45.4	88
v) blanks with compounds above acceptable limits?	40.9	59.1	88
vi) comparison of duplicate results outside acceptable limits?	36.4	63.6	88
vii) comparison of spiked and non-spiked samples outside acceptable limits?	50.0	50.0	72
viii) other:________________			
2) Is there a written record of the data manipulations performed by the system?	57.1	42.9	84
3) Is this record sufficiently complete to manually duplicate the data manipulations performed by the system?	70.6	29.4	68
4) In the data reviewed, were the data manipulations performed by the system found to be correct?	100	0	68
5) In the data reviewed, were quality control flags set by the system found to be in agreement with the original results?	93.8	6.2	64

YES NO RESP

6) If flags are changed on the system, is there documentation kept (either on the system or on hard-copy records) of both the changed and unchanged flags? 33.3 66.7 72

7) Are the flags of sufficient detail to characterize problems with the data (i.e., a flag merely setting the sample as invalid without providing detail as to the nature of the problem may not be sufficient)? 55.0 45.0 80

8) In those cases where data manipulations are not made by the system, was the information stored in the system an accurate reflection of the raw data? 90.0 10.0 80
If no, describe any problems encountered:

9) Are technical records maintained on the data system sufficiently complete as to allow scientific review of the data? 59.1 40.9 88

10) Are system maintenance records of sufficient detail and organization to determine when and what kind of maintenance was performed on the system? 60.9 39.1 92

11) Are records concerning release of new software versions of sufficient detail and organization to determine under what version all data was processed? 333.3 66.7 84

12) Are acceptance test records of sufficient detail and organization to determine what tests were conducted and the results of those tests? 47.6 52.4 84

VII

REFERENCES

1. Federal Register. Toxic Substances Control Act (TSCA); Good Laboratory Practice Standards. 40 CFR Part 792. Vol 52, No. 248, December 28, 1987, pp.48933-46.

2. Federal Register. Federal Insecticide, Fungicide and Rodenticide Act (FIFRA); Good Laboratory Practice Standards. 40 CFR Part 160. Vol 52, No. 248, December 28, 1987, pp. 48920-33.

3. Federal Register. Food and Drug Administration (FDA); Good Laboratory Practice Regulations. 21 CFR Part 58. Vol 52, No. 172, September 14, 1987, pp. 33763-82.

4. Computer Security Act of 1987. Public Law 100-235, January 8, 1988.

5. Environmental Protection Agency (EPA) System Design and Development Guidance. OIRM 87-02, 10/87.

6. Data Standards for the Electronic Transmission of Laboratory Measurement Results. EPA Order 2180.2, December 10, 1987

APPENDIX B

Description of Laboratory Systems and Personnel Structure

APPENDIX B(1)

No.	System	Responsible Person	Quality Assurance	Supervisor
1	Digital Microvax 2	Computer Specialist	Chief Organic Chemistry Section	Chief Analytical Support Branch
2	Concurrent SP3280	Chemist/LIMS Site Mgr./LAN Administrator	Environmental Scientist/Quality Control Officer	Quality Assurance Officer
3	IBM 4381	Sr. Systems Analyst	Q/C Coordinator	Chief Technical Support Branch
4	Encotec with IBM Clone (AT)	GC/MS Sr. Chemist/Systems Programmer	QC/MS Group Leader	Lab Manager
5	IBM Clones	President	President	Owners
6	Perkin Elmer LIMS 2000	Data Systems Manager	Quality Assurance Manager	Vice President
7	WANG VS-7110	Corporate Director of Information Systems	QA Department Manager	Sampling and Analytical Services Division Director
8	VG Sample Manager on a DEC Microvax 3600	Manager, Data Management	Manager, Quality Control	Lab Director
9	Concurrent 3230	System Analyst	QA Officer	Director
10	PC Network with Radian SAM v3.3	Technical Support	Quality Control Coordinator	Lab Director
11	Digital VAX 8650	Coordinator-Computer Systems Operations	Section Supervisor QA	Assistant Lab Director
12	Radian-SAM/LIMS on PC	Chief Chemist	QC Officer	Lab Director
13	Radian SAM v3.5 plus many customizations	Computer Systems Manager	QA/QC Coordinator	Lab Director

APPENDIX B(2)

No.	System	Responsible Person	Quality Assurance	Supervisor
14	Concurrent 3212	Systems Manager	Quality Assurance Coordinator	Vice President and General Manager
15(1)	HP 3000	Manager, System Development	Vice President, Quality Assurance	CEO & President
15(2)	HP 1000 interfaced to EPSON IBM	Section Chief	Chemist	Section Chief
15(3)	HP 1000 E-Series	Manager of Lab Automation	Vice President, Quality Assurance	CEO & President
15(4)	Compaq and Epson AT Compatible PC	Chemist	Chemist	CEO & President
15(5)	DEC Micro PDP-11/73	Sr. Systems Analyst	Not Provided	Not Provided
16(1)	IBM PC AT	Chemist	Compliance Director	Lab Manager
16(2)	HP Vectra	Management Information Director	Compliance Director	Lab Manager
16(3)	HP-1000 Varian	Chemist	Compliance Director	Lab Manager
16(4)	IBM or IBM Compatible PC/XT, AT	Programmer	Compliance Director	Lab Manager
17	Digital	Manager of Computer Services	Manager of Quality Programs	Director
18	Perkin Elmer LIMS	Site Manager, Sr. Lab Automation Specialist	Lab Director, Section Chief, Chemist	Lab Director