SCIEX
The Power of Precision

DATA INTEGRITY IN THE ANALYTICAL LAB

Author: Blair C. James

INTRODUCTION

Assurance of data integrity, security, and privacy, as required by regulators and essential for industry, is becoming more difficult for today's bioanalytical laboratory. This challenge is compounded by the growing size and complexity of typical datasets, which have expanded to include multiple analytical techniques, and span geographies, regulatory frameworks, and business models.

It is more important than ever that laboratory and information technology managers be proactive to organize, secure, and protect their data. This white paper discusses the issues related to data integrity and security when using typical computerized analytical systems. Although this white paper will focus on bioanalytical and clinical laboratories, similar considerations apply to food testing, environmental testing, forensic, and other laboratories.

QUALITY

Quality has been defined as:
"The degree to which a product meets requirements."

The product of bioanalytical, clinical, and other laboratories is data obtained by the analysis of physical samples. These data are used in support of decisions which can have profound effects on drug safety and effectiveness, patient treatment, environmental and food safety, public health, and law enforcement. If the data proceeding from a laboratory is not of high quality, the negative consequences can be as severe as death and loss of freedom.

When the quality of the data produced by the laboratory is poor, the negative consequences to the businesses and agencies can be severe, including financial losses, reputational damage, legal liability, and in the case of willful falsification of data, possible criminal prosecution and imprisonment.

The requirements that the data are intended to meet are superficially straightforward (e.g., that the compounds identified in a sample are correct, or that the concentrations reported are accurate.) Though when the consumers (i.e., customers) of the data produced by the laboratory are considered, the requirements become more complex.

LABORATORY STAKEHOLDERS

Business Stakeholders

The primary customers of bioanalytical and clinical laboratories are the principle investigators in the case of bioanalytical studies, and healthcare providers (and ultimately patients) in the case of clinical laboratories. These customers require that data are timely, complete, and accurate. The following table identifies these and other customers and stakeholders of the laboratory, and some of their needs and requirements.

Laboratory Customers and Stakeholders

Stakeholder	Description	Requirements and Concerns
Principle Investigators	The scientists intending to submit applications to regulatory agencies for approval of drug products.	• Accuracy, consistency, timeliness of data • Access to data throughout the study period • Effective analysis and reporting tools • Patient and participant privacy
Healthcare Providers	Clinicians seeking to diagnose disease and to monitor administration of drugs.	• Accuracy, consistency, and timeliness of data • Prompt and concise reporting of results • Patient and participant privacy
Study Sponsors	The business or/and government entities commissioning studies.	• Accuracy, consistency, timeliness of data • Access to data throughout the study period • Safekeeping and ready retrieval of data throughout the records retention period • Patient and participant privacy
Regulators	Government agencies and professional bodies that oversee bioanalytical or/ and clinical testing laboratories.	• Accuracy, consistency, timeliness of data • Ability to reconstruct study activities, results, and reports throughout the records retention period • Ability to identify altered or invalid data • Evidence of suitability of personnel, equipment, and processes for intended use • Patient and participant privacy • Adherence to the applicable regulations
Laboratory Managers	Individuals and groups responsible for the day-to-day operation of the laboratory	• Laboratory productivity • Reliability of personnel, equipment, and processes • Security and safeguarding of data • Patient and participant privacy
Information Technology Managers/ Departments	The individuals and departments charged with managing information resources and infrastructure in an organization.	• Reliability of equipment, including information technology infrastructure • Usability of equipment and software • Robust design of systems and processes to maximize uptime and productivity • System security • Patient and participant privacy
Quality Assurance Unit	The individuals and departments responsible for assuring that quality procedures are followed, regulatory requirements are met, proper records are kept, etc.	• Availability of records throughout the record retention period • Adherence to policies and procedures by organization staff • Proper documentation and justification for deviations from policies and procedures

A review of the customer and stakeholder needs in the foregoing table reveals common requirements:

- Efficiency and productivity
- Reliability; suitability for intended use
- Accuracy, consistency, and timeliness
- Short- and long-term protection of data
- Patient and participant privacy

REGULATORS

Governments and non-governmental organizations (NGOs), such as professional societies, standards bodies, etc., are all concerned with the foregoing requirements. In addition, these regulatory bodies are also concerned with the laboratory's ability to objectively demonstrate compliance with their regulations.

These regulators have expressed their expectation in the form of formal requirements and standards. Unfortunately, it is not possible to draft legal requirements and standards in enough detail to remove the need for interpretation by industry, even when the published regulations and standards are extensive. Table 2 lists a sampling of the many regulatory and professional organizations and their regulations:

Incomplete Listing of Regulatory Agencies and Regulations

Agency/Organization	Regulation
United States (Food and Drug Administration (FDA), US Department of Health and Human Services (HHS), Centers for Medicare and Medicaid Services (CMS), Centers for Disease Control (CDC))	21 CFR Part 58- Good Laboratory Practice for Nonclinical Laboratory Studies
	21 CFR Part 210, 211-Current Good Manufacturing Practice for Finished Pharmaceuticals
	21 CFR Part 11- Electronic Records, Electronic Signatures
	21 CFR Part 820- Quality System Regulation
	Data Integrity and Compliance With cGMP Guidance for Industry
	Health Insurance Portability and Accountability Act of 1996 (HIPPA)
	42 CFR Part 493- Clinical Laboratory Improvement Amendments
European Union (European Medicines Agency (EMA), European Commission)	Directive 2004/9/EC- the inspection and verification of good laboratory practice (GLP)
	Directive 2004/10/EC-the harmonisation of laws, regulations and administrative provisions relating to the application of the principles of good laboratory practice and the verification of their applications for tests on chemical substances
	Directive 91/356/EEC, as amended by Directive 2003/94/EC, and 91/412/EEC-Rules governing medicinal products in the European Union
	Annex 11: Computerized Systems
	Regulation (EU) 2016/679 (GDPR)-the protection of natural persons with regard to the processing of personal data and on the free movement of such data
	General European OMCL Network (GEON) Quality Management Document PA/PH/OMCL (08) 69 R7: Validation of Computerised Systems Core Document
International Society of Pharmaceutical Engineers (ISPE)	Good Automated Manufacturing Practice (GAMP) Records and Data Integrity Guide
College of American Pathologists	Standards for Laboratory Accreditation
Clinical Laboratory Standards Institute (CLSI)	Laboratory Instrument Implementation, Verification, and Maintenance (GP31)
China Food and Drug Administration (CFDA)	Good Laboratory Practices for Non-Clinical Studies of Pharmaceuticals, 2003
Organisation for Economic Cooperation and Development (OECD)	OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring; OECD Principles on Good Laboratory Practice (ENV/MC/CHEM (98)17)
	OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring; Application of GLP Principles to Computerised Systems (ENV/JM/MONO (2016)13)
Japanese Ministry of Health	Ordinance on the GLP Standard for Conduct of Nonclinical Safety Studies of Drugs
World Health Organization (WHO)	Handbook: Good Laboratory Practice (GLP): Quality Practices for Regulated Non-Clinical Research and Development - 2nd ed.
International Organization for Standardization (ISO)	ISO/IEC 17025:2005-General requirements for the competence of testing and calibration laboratories ISO 15189:2012-Medical laboratories – Requirements for quality and competence
U.S. Pharmacopeia	USP 1058-Analytical Instrument Qualification
Pharmaceutical Inspection Co-operation Scheme (PIC/S)	Good Practices for Computerized Systems In Regulated "GXP" Environments

Regulators expect to find evidence that individuals and departments involved in the production of data are in control of the critical processes that have the potential to impact the quality of those data. This expectation extends to the suitability of personnel, processes, equipment, and facilities for the use(s) for which they are intended.

The balance of this whitepaper will describe an approach to satisfying regulators and other stakeholders as to the integrity of the data produced by a laboratory when using computerized analytical systems.

DATA INTEGRITY

Introduction

Data Integrity is defined as:
"The degree to which a collection of data is complete, consistent, and accurate."

Completeness: Data are complete when all the information required to support a decision is present and accessible. Good science requires that all measurements and calculations can, at least in principle, be reproduced by scientists other than those who performed the original experiment. Therefore, a set of data cannot be considered complete unless all the information required to reproduce the experiment and analysis are present.

Consistency: Data are consistent to the extent that they support (i.e., do not contradict or undermine) other observations or datasets.

Accuracy: Accuracy indicates that the individual observations in the dataset, and the dataset itself expresses factually correct information. Individual observations must be both precise (i.e., repetitive measurement of the same item return comparable values), and accurate (i.e., each measurement reflects a factually correct value).

Data Lifecycle

A data item represents a measured value at a certain point in time, under specific conditions. A data set is a collection of data items. These simple definitions fail to capture the complex processes that result in data being useful and reliable to any organization.

The ISPE GAMP Guide suggests the following data lifecycle:

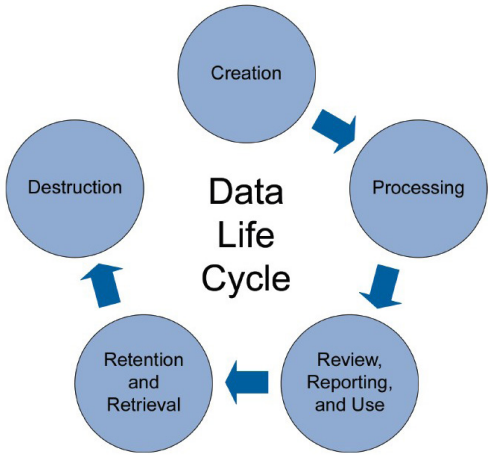

Figure 1 Data Life Cycle.
Source: ISPE GAMP Records and Data Integrity Guide

Data integrity is ensured following correct processes at each phase of the data's lifecycle. Issues to consider at each phase include:

1. Creation Phase
    - Selection of appropriate instrumentation and software
    - Proper installation, qualification, and maintenance of instrumentation and validation of software
    - Appropriate analytical methods, including use of quality control samples
    - Correct use of the analytical instrumentation and execution of the analytical method
    - Capture and retention of all necessary data and meta-data
2. Processing Phase
    - Selection of appropriate processing methods which are robust enough to need little or no human-assisted processing
    - Identification of iterative processing techniques (reprocessing), and retention of intermediate processing states
3. Review, Reporting, and Use
    - Written procedures for data review and approval, including consideration of:
        - Completeness of meta-data produced in prior phases
        - Use of human-assisted processing
        - Reprocessing
        - Audit trails
        - Detection and avoidance of testing into compliance
    - Written procedures for reporting, including:
        - Accuracy and security of reports
        - Ability of users to determine which results are reported
        - Ability of users to alter presentation, including scaling, labeling, excluding observations etc.
        - Appropriate review and approval of reports
    - Use:
        - Reports distributed to appropriate individuals
        - Appropriate level of review and approval for data used for decision-making
4. Retention and Retrieval
    - Records retention period mandated by regulatory bodies
    - Records retention period based on user and business needs
    - Ability to view and re-process data at a future date
    - Archive growth over time
    - Redundancy of archived records
    - Protection of the archive, e.g.,
        - Usable life of storage media
        - External and environmental threats to the archive
5. Destruction
    - Written policies and procedures for identifying records subject to destruction
    - Record-keeping of data destruction activities

CONTROLS

Regulators expect laboratory management to be in control of the critical processes that have the potential to impact data quality.

Procedural Controls vs. Technical Controls

A procedural control consists of the design and implementation of a process that, when followed, ensures data quality and integrity. Standard Operating Procedures and Work Instructions are two common types of procedural controls in the laboratory. It is important that to the extent possible, procedures are designed to be self-enforcing and self-correcting, and procedures should anticipate known risks to data integrity and mitigate them.

A very simple example of a procedure control is a sign on an unlocked door that states "Authorized Personnel Only". So long as individuals who may want to go through door correctly know whether they are "authorized", and so long as they choose to follow the procedure, the room will remain secure. As soon as someone mistakenly thinks that they are authorized, when in fact they or not, or they simply decide to violate the procedure, the procedure fails.

Technical controls are generally the features of hardware, software, premises, and other infrastructure that limit the way in which systems can be used. Technical controls can be more effective than procedural controls, because the system itself limits the ability of individual users to violate the intent of the control.

An example of a technical control is the installation of a lock on the above-mentioned door. Now, independent of individual choice, personnel will only be able to gain access to the room if they have the correct key.

Technical and procedural controls are usually used in tandem. For example, to secure a room, the organization might install a lock on the door and write a procedure specifying the process for requesting access and evaluating those requests to determine to whom a key should be issued.

If a technical control is available, it must be properly configured and used. A procedural control alone is never adequate when a technical control can be implemented.

ENSURING DATA INTEGRITY OF ANALYTICAL SYSTEMS

This section describes actions that must be taken to ensure data integrity of a computerized analytical system.

Premises and Infrastructure

1. Security

Access to the laboratory and supporting facilities must be restricted to those personnel having a legitimate need and with the appropriate training. Access records must be maintained, and visitors must be positively identified. Some organizations, and particularly those handling dangerous or bioactive substances, should provide training and obtain positive consent from individuals accessing the areas where they may be exposed to danger. Personnel must be provided with appropriate personal protective equipment.

2. Size and Capacity

Facilities, including storage, and instrumentation must be of adequate size and capacity for the work performed. Benchtops must be large enough to perform work safely and must be kept clean to prevent hazards and the possibility of cross-contamination of samples and reagents.

3. Training and Competency

Personnel performing work in the laboratory must have the education, training, and experience for the work performed. Records of education and training must be maintained and must be available for inspection by regulators. Training should include relevant external regulations, organizational policies, standard operating procedures, and safety and environmental standards.

4. Vendor and Instrument Selection

Systems, including analytical instruments and any supporting computer system(s), must be suitable for their intended use. Vendors should be qualified to ensure that they have an adequate quality management system including internal policies and procedures. Standards certifications, such as ISO 9001 certification are useful indicators of the vendor's internal controls. Vendor audits must be performed to document relevant vendor characteristics.

5. Instrument Installation

Instrumentation in the laboratory must be installed as specified by the vendor with attention to:

- The physical environment must be suitable for the instrument, including temperature, humidity, electrical supply, gas supply, venting, waste disposal, etc. Computers used to operate and support the instrument require these, as well an appropriate security and networking environment.
- During installation, information technology requirements must be addressed, such as compatible versions of firmware and application software. Backup and disaster recovery procedures should be implemented as soon as the system installed, so that effort invested in configuration and testing can be preserved should a system failure occur.

6. Hardware Qualification

Instrument hardware must be qualified to demonstrate that the instrument operates and performs as intended by the vendor. Hardware qualification should be performed according to vendor specifications to ensure that the hardware, computer systems, and software are installed and configured properly. Like the individuals that use the instrumentation, hardware qualification should be performed by personnel that have the education, training, and experience for the work performed. Records of the personal training of the qualification personnel must be maintained with the qualification records. Re-qualification must be performed periodically, preferably annually or bi-annually, and after any change having the potential to impact system operation and performance.

When an instrument is relocated, the entire qualification procedure, including installation, operational, and performance qualification must be performed.

Hardware qualification should include:

- Installation Qualification: To verify that the system is installed as specified, consists of the proper components, and to document that the installation has been performed in a controlled manner.
- Operational Qualification: To confirm that the system operates according to vendor specifications. Operational Qualification may or may not demonstrate the suitability of the system for its intended use, depending on the vendor specifications verified and the ultimate application of the system. Following preventative maintenance or repair, critical operational parameters must be confirmed by re-execution of the operational qualification. Routine analytical tests do not satisfy the requirements for qualification. The qualification should include traceability of all standards and certificates for calibrated tools used in execution of the qualification.
- Performance Qualification: To demonstrate and document that the individual components of the system function together in a manner comparable to a quantitative assay under near-real-world conditions.

7. Software Configuration and Validation

Most analytical systems rely on computers and computer software to acquire, evaluate, analyze, and report the data produced by the system. It is critical that computerized systems be validated to ensure consistent intended operation and suitability for purpose.

US FDA provides guidance for software validation in General Principles of Software Validation. This guidance applies specifically to medical device software and 21 CFR Part 820, Quality System Regulation, but the approach described is based on well-established software development principles, and therefore can be applied to any software.

The software validation is defined as "confirmation by examination and provision of objective evidence that software specifications conform to user needs and intended uses, and that the particular requirements implemented through software can be consistently fulfilled."

The International Society of Pharmaceutical Engineers (ISPE) has provided guidance and methodology for ensuring the integrity of data in the Good Automated Manufacturing Principles (GAMP) 5 Guide: Compliant GxP Computerized Systems. Because it is not possible to test every aspect of a software application under all possible scenarios (due to time and resource constraints), GAMP 5 describes a risk-based approach to software validation, in which the risks to data integrity are evaluated and prioritized, so as to make best use of limited testing resources, and to guide the investment of validation resources in activities appropriate to the nature and intended use of the software. One common approach to risk analysis is to assign each identified potential risk a rating based on the likelihood of the risk occurring, the impact to data quality should the risk materialize, and the likelihood of the occurrence being detected during normal operations. These three factors can be assigned a numeric rating, and then combined

See US Pharmacopoeia 37, <1058> Analytical Instrument Qualification, for additional guidance for instrument qualification.

arithmetically using a weighting strategy to arrive at a risk priority. This allows available resources to be applied to mitigate the most important risks. See US FDA General Principles of Software Validation; Final Guidance for Industry and FDA Staff, January 11, 2002 for more information.

Throughout the validation process, artifacts will be produced to document the intent, execution, and result of validation. These include:

- Validation Plan: Defines the participants, roles and responsibilities, documentation and activities required for the specification, development, testing, review, and release of the computerized system into the regulated laboratory environment.
- Risk Assessment: Identifies and prioritizes risks to data integrity and identifies strategies to mitigate those risks.
- User Requirements Specification: The user, functional, security, and electronic records requirements of the computerized system intended for use in the regulated laboratory environment.
- Configuration Design Specification: Records the configuration settings required to ensure data integrity and fulfillment of user requirements.
- Test Plan: Describes the strategy and approach to testing the computerized system.
- Installation, Operational, and Performance Qualification test scripts: Contain the specific test procedures used to verify that requirements are satisfied.
- Traceability Matrix: Links (traces) requirements to specific tests to ensure that no requirements remain unaddressed and to provide a convenient reference to the tests performed to verify that the requirements have been fulfilled.
- Validation Summary Report: Summarizes the validation activities performed and the outcome of those activities. Approval of the Validation Summary Report authorizes the release of the system for regulated use.
- Regulatory Compliance Assessment: A list of the specific regulatory requirements addressed during validation, and a description of how the validated system, including hardware, software, and procedures, satisfies those requirements.

8. Data Security and Privacy

Data must be secure and kept private throughout the analytical system's life cycle. Most analytical software provides functions for limiting access to the system and to specific functions of the system. These roles and privileges must correspond to each user's job responsibilities, while considering the education, training, and experience of the user.

Many global regulatory regimes require that systems that create electronic records include an automated audit trail to record the actions taken by users while operating the system, including the authorship, modification (including deletion), and approval of data. To the extent that the system allows configuration of the audit trail, it must be configured so that actions that have the potential to affect data quality are captured with required details, such as the user performing the action, the date and time, and the existing data values prior to changes.

Many governments and organizations require that the privacy of data be protected. If a system stores personal data, the system must provide the functions to protect the privacy from disclosure or misuse.

Every computerized system must have a system administrator who can make potentially destructive changes to the data and configuration. This is the nature of computerized systems. This role must never be assigned to an individual who uses the system to generate, process, or review data. Many laboratories assign this role to the information technology department. All laboratories should create and follow procedures to control how these operations are requested, evaluated, approved, and performed.

9. Data Processing

Data from more-complex analytical systems are generally not useful in their raw form. For example, the raw time and intensity data acquired from a LC-MS system look like a very long series of numbers and must be quantified to calculate the concentrations of analyte in the sample.

Regulators expect the quantitation parameters (also called methods) to be robust enough that only minimal adjustments are required to produce peak areas, calibration curves, etc. If a run of samples is acquired and processed, and the quality control samples in the run fail, the user might be tempted to adjust the quantitation parameters repetitively until the QCs pass, then report the results of the run using those successful parameters. This practice is sometimes referred to as "testing into compliance". Such practices undermine the scientific basis of the analysis. A similar situation arises when chromatographic peaks are integrated manually.

It is critical that systems be configured so that re-processing of data, as described above, is prevented, and that attempts to do so are apparent in the audit trails. Manual integration must be controlled so that its misuse is prevented.

10. Data Review and Approval

The Quality Assurance department must have written policies and procedures for the review of data. QA personnel should be trained to be able to detect errors and misuse of processing tools when reviewing data. Data review should include audit trails, and written procedures should identify how to detect and investigate occurrences of re-processing and manual integration.

11. Records Retention and Archival

Archival in the context of GxP regulations is the act of preventing users from having access and making changes to study data after the acquisition and analysis phase of the study is completed. For example, records might be archived by copying the records to a secure network storage location accessible only to the Archivist (the individual(s) that have access to and responsibility for the data after the data are archived). The Archivist should not be a member of the team participating in the study. Many organizations delegate this responsibility to the Information Technology department. Procedures must be implemented for the archival of the data at the appropriate time, handling of requests to extract data from the archive, and protection of the data throughout the records retention period.

Retention periods vary by agency, region, country, etc. Both regulatory and business needs must be evaluated when determining the retention period for a given type of record. Most organizations are best served by considering the retention period to be infinite and planning the archival and retention processes accordingly.

Routine backups of data do not satisfy archival requirements, because routine backups do not include provisions for restrictions on access, retrieval of data from the archive, and organization by study. This is not to say that the archive cannot be protected by backup procedures, only that the specific archival requirements are usually not fully satisfied by routine backup procedures.

12. Backup and Disaster Recovery

Data generated by the analytical system must be protected from loss for business and regulatory reasons. It is expensive and difficult, often impossible, to reacquire sample data lost through equipment failure, environmental disaster, or human error. Information Technology industry best practices for data backup are beyond the scope of the paper but some principles can be stated to guide the planning of a disaster recovery process:

- Data backup should be automated to reduce the probability and impact of human error.
- Data should be copied to backup as soon after the data are generated as possible.
- Backup copies of data should be verified against the original copy immediately after creation.
- Copies of the data should be maintained on-site for ready retrieval, as well as in off-site secure storage to mitigate the effects of local environmental disasters, such as flood or fire.
- A backup routine that rotates media and makes multiple copies of data on multiple media should be used so that the backup copies are redundant. For example, on Friday a full copy of all data, with incremental backups of the data generated each day on other days of the week. On the next Friday, an additional complete copy of all data is made, and so on.

13. Ongoing Activities

After an analytical system has been installed, qualified, and the software validated, users must perform periodic activities to ensure that the system remains in the validated state. Otherwise, the system may become unsuitable for its intended purpose through ordinary use and system changes.

i. Change Control

Once a system has been installed and validated, subsequent changes must be controlled to maintain the validated state of the system. A full discussion of change control strategies is beyond the scope of this paper, but any change control process should include the following steps:

- A process for requesting changes.
- Evaluation of proposed changes to identify the risk the change presents to the integrity of data.
- An appropriate level of testing to mitigate the risks identified.
- Assessment of the need for data conversion or migration.
- Documentation of the change.
- Ongoing monitoring of the quality of data.

ii. Preventive Maintenance

All analytical systems must be maintained per the specifications and schedule recommended by the vendor. This may include cleaning, tuning, calibration, etc. It is critical that these activities be performed on time and properly documented.

iii. Periodic Requalification

Systems must be requalified, including OQ and PQ, periodically to ensure that the quality of the data produced remains high. It is convenient to re-qualify after preventive maintenance, which ensures that the operational and performance characteristics of the system are within vendor specifications after the maintenance is performed. Analytical systems should be requalified at least annually, or more frequently in high-throughput environments.

Requalification must also be performed after any repair that may impact system performance.

iv. Software Configuration Maintenance

Software validation activities establish a baseline of the system configuration. To ensure that this validated configuration is maintained, systems should be audited annually. The audit should include verification that configurable settings agree with the validation baseline, that un-needed or un-used user accounts are deactivated, and that procedures for use of the system are effective to ensure high data quality and integrity.

CONCLUSION

Organizations make a huge investment in personnel, equipment, software, sample collection, and analysis. Deficiencies in data integrity can result in loss of this investment, and in the case of severe regulatory gaps, even the viability of the business. By implementing policies and procedures to ensure data integrity and regulatory compliance, organizations can minimize these risks while still operating efficiently.

AB Sciex is doing business as SCIEX. © 2019 AB Sciex. For Research Use Only. Not for use in diagnostic procedures. The trademarks mentioned herein are the property of AB Sciex Pte. Ltd. or their respective owners. AB SCIEX™ is being used under license.

RUO-MKT-19-9543 (05/2019)

Headquarters
500 Old Connecticut Path
Framingham, MA 01701 USA
Phone 508-383-7700
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices

SCIEX
The Power of Precision