Journal of the Society for Clinical Data Management

Laboratory Data Handling. *Journal of the Society for Clinical Data Management*. 2023; 1(1): 15, pp. 1–8. DOI: https://doi.org/10.47912/jscdm.335

GCDMP

REVIEW ARTICLE

Laboratory Data Handling

The vast majority of clinical studies use laboratory data, which should be treated with the same rigorous attention to detail and data quality as any other clinical data. This chapter describes different types of laboratories, different types of laboratory data, and important elements of laboratory data handling. In particular, the chapter discusses the importance of standards and reference ranges for laboratory data, as well as principles and processes to help ensure the accuracy and integrity of all laboratory data.

Keywords: Clinical Data Management; Local Lab, Central Lab; Lab Data Management; Good Clinical Practice

Introduction

The word "lab" (or "laboratory"), is defined by Merriam-Webster as "A place equipped for experimental study in a science or for testing and analysis." Within the context of clinical data management (CDM), labs are where biologic samples such as blood or urine are sent for analysis, or diagnostic images or data such as electrocardiograms or Holter monitors are evaluated or interpreted. Because the results of these tests do not originate from a case report form (CRF) at a study site, these types of external data are often transferred as electronic files.

Lab data are used in most preregistration clinical studies and proper handling of these data is crucial to the success of a study. CDM personnel are responsible for data integrity throughout all lab data transfer and cleaning activities. CDM personnel may also be involved with setting up standards and processes for their organization to help ensure the integrity of all data, including those from labs.

Scope

This chapter describes differences between various types of labs and lab data, as well as how CDM practices may vary in different situations. For the purposes of this chapter, the term "lab" generally refers to lab vendors, as opposed to lab tests, which will be referred to as "tests" or "lab tests." Although local and central labs are not the only lab types discussed, the distinctions between local and central labs can also apply to specialty labs, core labs, and virtual central labs. Specialty labs and core labs may operate as either central or local labs, while virtual central labs operate as central labs.

Also, most CDM processes relating to lab data handling primarily vary between local and central labs. As such, the main focus of this chapter will be on local and central lab data handling.

Some of the tasks described in this chapter may be joint responsibilities between different groups, just as there may be many different groups involved in the implementation of various tasks. However, clinical data managers need to be conscious of whether or not these tasks have been performed in a satisfactory manner.

Minimum Standards

- Maintain standard operating procedures (SOPs) for all processes relating to lab data collection, transfer, and validation of data loading and data feasibility.
- Identify labs involved with a study as early in study setup as possible.
- Use standardized names for lab tests and units.
- Ensure reference ranges are defined prior to first data receipt when using a central lab.
- Where possible, ensure reference ranges are defined prior to first data receipt when using a local lab.
- Ensure updates to reference ranges are obtained and implemented in a timely fashion.
- Document all data transfer specifications thoroughly when using labs transferring data electronically.
- Determine software/hardware required to access data prior to a test transfer and ensure the format of the data medium is compatible.

Best Practices

- Use accepted standards such as those from Clinical Data Interchange Standards Consortium (CDISC) when possible.
- Define all lab data standards prior to beginning data collection.
- Ensure reference ranges are defined for population subgroups (e.g., ethnicity) that differ significantly from other defined groups or subgroups.
- Implement a standard process to collect and archive reference range data.
- Use a standard method of data review for local lab data and reconciliation of central lab data.
- Develop a data transfer agreement for electronic transfers and perform quality control of the test transfer.
- Document and confirm all lab variables prior to signing off on data transfer specifications.
- Implement a conversion factor table to standardize conversion of conventional units to the International System of Units (SI).
- Define edit checks for inclusion/exclusion criteria based on lab data and route to appropriate team members to review.

- Use standardized units so that performing edit checks on converted data produces a more consistent review of results.
- Send requests for central lab data corrections using a formalized process, for example, on a correction log sent to the lab vendor to update and return after correcting and resubmitting the lab data file.
- Implement a system to manage data collected outside protocol parameters.

Distinctions Between Types of Labs

Although data managers most frequently work with central labs or local labs, other kinds of labs include virtual central labs, specialty labs and core labs.

These types of labs tend to fall under the categories of local or central in regard to many processes and characteristics. This section details each type of lab and defines which tests and processes they support. **Table 1** details advantages and disadvantages of each type of lab. Advantages and disadvantages may vary geographically, due to regional variations in definitions of various types of labs.

Central Labs

A central lab processes lab samples from multiple clinical sites or studies at one central location. These labs often support multicenter and international studies. Central labs can process many types of samples but most commonly process and report clinical chemistry, hematology and urinalysis. Central lab data are typically transferred electronically from the lab to the sponsor or contract research organization (CRO) throughout the course of a study, resulting in rapid and continuous data transfers and improved safety review and study management. Most central labs have their own file formats but are willing to work with sponsors or CROs at the beginning of a project to define data transfer specifications. Establishing these specifications up front streamlines the process of data transfers.

Local Labs

Local labs are labs in close proximity to individual clinical study sites or patients and are most often used when timely results are needed. Local labs may also be known as "regional" or "preidentified" labs in some locations, such as parts of Europe. Local labs are commonly used in oncology studies, where lab results could be the deciding factor on dosing or not dosing a subject. Each local lab must provide a set of reference ranges to the sponsor or CRO, which increases the work needed for all aspects of lab data collection and integration with study databases. Local labs are typically not able to perform electronic data transfers, so sites become responsible for entering this information onto CRFs. This process can be very time-consuming and error prone, resulting in an increase in the number of queries to the site for clarification or correction.

Table 1: Advantages and Disadvantages of Lab Types.

Type of Lab	Advantages	Disadvantages
Central Labs	Uses one set of analytical equipment, methodologies, kits and reagents Provides training and instructions for collection and shipping of samples, as well as safety alert notifications Standardized results from one set of reference ranges and units Access to lab results in near real time once samples are received and analyzed	Logistical support and costs for shipping lab samples The turnaround time needed to receive central lab data may be too long when immediate results are needed
Local Labs	Lower costs and shorter turnaround time due to not having to ship samples Local lab experience with processing samples from their subject population	Greater potential for errors due to paper-based data transfers and differences between reference ranges from one lab to another Variability in the methods used to perform tests Variability in reference ranges and units used for measurement Reference ranges may be more difficult to obtain
Virtual Central Labs	Reduced shipping costs Decreased need for resampling due to samples becoming compromised during shipment Simpler data processing due to having a central calibrator	Requires detailed process and quality control (QC) measures to ensure lab results are reproducible with minimal variance from site to site
Specialty Labs	Highly experienced and qualified for performing specialty tests	Many specialty tests require more time to generate test results
Core Labs	More focused quality control, more accurate results and a higher degree of standardization and specialization within a designated area	Additional time may be incurred for centralized processing

Virtual Central Labs

The virtual central lab (VCL) is typically a group of labs located throughout the world that are under the umbrella of one company (or partnership). The VCL is based on a central calibrator that runs in parallel with lab samples from all labs participating in a clinical study. The calibrator and lab sample results are compared and results are adjusted based on the calibrated value used by all the labs participating in the study. This process reduces the logistics and costs of shipping lab samples.

Specialty Labs

Specialty labs are used to analyze samples or run assays for nontraditional (or esoteric) tests, which are typically tests that take a considerable amount of time and effort to produce. The amount of time needed is typically outside the control of the lab, although the longer timeframe for these test results must be considered when planning a clinical study. Examples of these tests include biomarkers, genetic testing, pharmacokinetics, and isolation of cancer genes. Specialty tests may be conducted by one lab to ensure standardized results, which is vitally important because these test results are often used as primary efficacy variables.

Core Labs

For the purposes of this chapter, core labs are labs that specialize in a particular therapeutic area or body system. Examples of core labs include stem cell core labs, electrocardiogram (ECG) core labs, imaging core labs, cardiovascular core labs, hematology core labs and oncology core labs. Core labs are vitally important in large clinical studies for their accurate results, which may be used to interpret or support primary or secondary endpoints.

Lab Data in Clinical Studies

Lab data usually fall under the categories of safety, efficacy or specialty data. There are, however, instances where data may fall into more than one of these categories, such as efficacy data that also relate to a safety parameter.

- Safety Data—Lab data can be used to identify or quantify deleterious biological processes occurring in a subject. One of the main purposes of safety data is to provide a baseline at screening of a standard battery of tests that can be repeated during the study to ascertain if there are any detrimental changes to a single parameter or panel. Examples include cardiac biomarkers released into the blood when heart tissue is damaged, or glucose levels in a diabetic population. Many lab tests performed in preregistration studies are performed for safety testing. These tests provide data for a warning system to detect potential safety concerns before they are observable as signs or symptoms.
- Efficacy Data—Efficacy data are typically lab data relating directly to the effectiveness of the study treatment. For example, in a study of a new drug intended to battle high cholesterol, one of the primary measures would be lab results of the subject's cholesterol levels in the bloodstream.
- Specialty Data—Specialty data may consist of genomic, proteomic or pharmacokinetic data from a specialty lab. These data do not always relate directly to safety or efficacy, but may be very informative with regard to underlying biologic or genetic processes. The following types of data are those most commonly collected by specialty labs.
  - Genomic—Genomics is the study of the genes of an individual at the DNA (genotype), mRNA (transcriptome) or protein (proteome) levels. Another variant of the study of genomic data is pharmacogenomics, which is the study of how an individual's genome affects the body's response to drugs. Pharmacogenomics may be instrumental in personalizing treatments for greater efficacy and safety.
  - Proteomic—Proteomics is the study of proteins produced by an organism or system, particularly the proteins' structures and functions. The proteome is the entire complement of proteins, including modifications made to a particular set of proteins. Proteomics is often considered the next step after genomics in the study of biologic systems. Proteomics, however, is much more complicated than genomics, because while an organism's genome is constant, the proteome differs from cell to cell and over time.[1]
  - Pharmacokinetic/pharmacodynamic—Pharmacokinetics studies drug absorption, distribution, metabolism, interaction and excretion. Drugs exist in a dynamic state within the body, and different drug events often occur simultaneously. To describe a complex biologic system, simplifying assumptions are often made concerning the movement of drugs. A pharmacokinetic model is conceived using mathematical terms, which are a concise means of expressing quantitative relationships. The intensity of the pharmacologic or toxic effect of a drug is often related to the concentration of the drug. For example, monitoring the concentration of drugs in the blood or plasma confirms that the calculated dose actually delivers the plasma level required for therapeutic effect. Pharmacokinetic models allow more accurate interpretation of the relationship between plasma drug levels and pharmacologic response.[2]
  - Biomarkers—Biomarkers are substances that are objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes or pharmacologic responses to a therapeutic intervention. According to some experts, to be defined as a viable biomarker, the biomarker should meet the following conditions:
    - Highly sensitive and specific in detecting a desired characteristic
    - Validated in postmortem confirmed cases
    - Standardized with sound bioinformatics
    - Specific for the desired characteristic compared with related disorders or biologic states

- Reliable in many testing environments and labs
- Minimally invasive
- Simple to perform
- Inexpensive[3]

Standards

The more standardized lab data are, the easier they will be to collect, process, combine, analyze and submit. Although standardization during study setup is optimal, standardization may also be performed during lab data collection or analysis of final results. A number of data standards have been published or are in development by Clinical Data Interchange Standards Consortium (CDISC), including a standard specific to lab data (LAB). For more information on CDISC standards, visit http://www.cdisc.org.

Test Names

Test names are the easiest and most common part of lab data to standardize. If using a central lab, a list of test names should be provided by the lab at the inception of the study. If using a local lab, test name standards would be applied when setting up the clinical database and CRF entry screens.

CDISC terminology for lab test names and test codes can be used to standardize results for a local or central lab. The CDISC controlled terminology model consists of an alphabetical accounting of the most common test names (long name) and test codes (short name). By utilizing CDISC controlled terminology for test names and test codes, sponsors and CROs can reap the benefits of less conversion time when preparing for submissions to regulatory bodies. In addition, if multiple studies are being conducted, the format of data has been established, and table templates and analysis dataset structure can be predefined and programmed earlier in the process.

Units

Although not encompassing all potential analytes, the most universal format to capture lab data is the International System of Units (SI). The following quotation describes SI units and the history of their development.

'SI units' is the abbreviation for le Systeme International d'Unites. These units are the result of over a century of international cooperation to develop a universally acceptable system of units of measurement. The SI is an outgrowth of the metric system that has been widely used throughout most of the world, but which has had little impact outside scientific fields in the United States, even though Congress passed the Metric Conversion Act in 1975, which endorsed the SI.

The SI is a uniform system of reporting numerical values permitting interchangeability of information between nations and between disciplines. The SI not only provides a coherent system of units, but also ensures that units are uniform in concept and style. A coherent system is one in which interconversions between the units for different properties requires the factor 1 only. With the SI, quantities can be more easily compared by means of the reduction in the number of multiples and submultiples in common use.[4]

SI units have almost complete worldwide acceptance and do not need any further conversion. In addition to an SI unit, most tests in the US are also associated with a conventional unit, which is typically based on US measuring methods. When lab test results are collected, the data must be standardized and converted to one common unit before analysis can begin. This can be a time-consuming task, especially when working with multiple local labs, each using a variety of conventional units.

One way to make unit conversion easier is to develop an internal conversion factor table using publicly available references. A table can be created for all tests, listing the most common conventional units as well as the conversion factor to transform to SI units. This conversion table will take significant effort up front; however once completed and verified it will save an enormous amount of time by being applied to subsequent studies.

Unexpected/Unscheduled Lab Data

During the course of a clinical study, lab tests are performed according to the schedule of the protocol. Sometimes an investigator decides to order a lab test outside protocol parameters, usually when a subject is experiencing adverse events or exhibiting symptoms of another disorder. When these lab tests are performed, they are considered unexpected or outside the protocol.

When the results of these tests are received from a central lab, they may be kept in a separate dataset from protocol-specified tests or flagged to ensure they are easily recognized and are not part of eventual study analyses. When these tests are received from a local lab, the CRF should be designed to capture results from these unexpected tests. Because these tests will not usually be known in advance, the CRF should be as generic as possible to accommodate study-specific variations, but should include the following fields.

- Test name
- Test result, reference range upper and lower limits, high and low values, and units
- Lab name
- Sample collection date
- Comments section for capturing why tests were ordered and to describe results of the tests

Unscheduled lab data, on the other hand, refers to tests that are within the scope of the protocol but are not performed according to the time and events schedule. This may occur for a number of reasons, including follow-up tests due to previous abnormal values, a subject's unavailability for sample collection at a specified time, or damaged samples (which may be classified as repeat lab tests by some organizations). These tests are captured in the same manner as scheduled sample collections, but must be identified as unscheduled data. For unscheduled results from a central lab, the lab should have a way to differentiate unscheduled sample collections from those that are scheduled. One convention is to have the visit number left blank and the visit name labeled as "Uns" or "U" for unscheduled, although some organizations may

design a numbering convention in advance for these circumstances. The sample collection date will then be used to sequence the sample collection among others for that subject. For local labs, the CRF should capture the lab name, sample collection date and unscheduled status.

Lab Reference Ranges

Lab results are of little value without the ability to analyze the results in comparison to other values. Lab results are typically either compared with other samples taken from the same subject at a different time point (e.g., baseline values), or are compared with a reference range. Reference ranges can also be known as "normal ranges," although not all populations can be considered truly "normal." Reference ranges are established by analyzing a large number of samples and statistically determining the appropriate reference range. Because values may differ according to variables such as age, gender, disease processes, or regional variations, multiple ranges are often established for a given test. Labs may either establish their own set of reference ranges or obtain ranges from published sources. Reference ranges typically consist of a high value, a low value, the unit of measurement, and an effective date. Reference ranges can also be age- and gender-specific, necessitating identification of these parameters. These values need to be collected only once per study unless there are changes to the specimen collection, instrumentation or methodology. Lab relicensure may also trigger the need to update documentation of reference ranges.

Use by Clinicians During a Study

In clinical studies physicians use lab results to determine if a subject meets study enrollment criteria and to monitor the subject's safety profile or efficacy effects, which may be attributable to the treatment received or from existing or new conditions. Physicians may use other tests to confirm a diagnosis or eliminate error due to false-positive results. They are aware that the reference range provided by a lab has confidence limits and that some normal individuals will have a value outside the reference range. Therefore, most physicians will consider a result normal if it is within the reference range, suspicious if it is slightly outside the range, and abnormal if it is considerably outside the range. Ultimately, the clinical assessment will determine if a particular analyte has clinical significance.

Use by Statisticians in Data Analysis

Biostatisticians view lab values through summaries of data, often comparing the proportion of subjects with out-of-range values to the proportion of subjects with values within the expected range. Biostatisticians also look at changes within subjects and summarize and compare those changes between treatment groups. Shift tables are used to present categories of test results before and after an action, such as study treatment, presenting classification comparisons such as "High-High," "High-Normal," "High-Low," "Low- Normal," "Normal-Normal," etc. Biostatisticians also use flags present in the lab data as cut points to identify out of range values, such as "H" for an abnormal high value or "C" for a critical value.

Collection of Reference Ranges

ICH 8.2.11 requires that "...normal value(s)/range(s) for medical/laboratory/technical procedures(s) and/or test(s) included in the protocol..." be located in the files of the investigator/institution and sponsor.[5] Also, the Clinical Laboratory Improvement Amendments (CLIA) require that labs have reference ranges for all test results produced. The collection of reference ranges is imperative to appropriately handling lab data.

Changes in Reference Ranges

ICH 8.3.7 requires that "Updates to normal value(s)/range(s) for medical laboratory/technical procedures(s)/test(s) included in the protocol..." be located in the files of the investigator/institution and sponsor.[6] Reference ranges are generally not changed or revised unless a new methodology is adopted, primary reagents are modified, or new instrumentation is introduced into the lab. Minor changes in the reference ranges of an analyte may not be significant due to the precision of the method. However, if there is a change in units or a large shift in the reference range, the new range should be used for any results after the effective date of the change. Changes to reference ranges and the effective date of the change(s) should be quickly communicated by the lab and/or investigator to the sponsor or CRO, and all changes should be clearly documented.

Importance of Population-specific Ranges

Many variables complicate establishing reference ranges, including sex, age, ethnicity, weight, geography, or time of specimen collection. Reference ranges should be defined for each subgroup that differs significantly from another subgroup. When ranges are not divided into subgroups, there may be a broadening of the reference range and loss of discriminatory power.

Variations in reference ranges are most commonly seen between different sex and age groups.

Lab Processes in Studies

Local Labs

When using local labs, more responsibility is placed on the site to record information. The process begins with obtaining and identifying a sample, then sending it to the local lab for analysis. Once the sample is tested and the report is received at the site, it is the responsibility of the primary investigator or subinvestigator to assess the lab report and determine if out-of-range values are deemed clinically significant (CS) or not clinically significant (NCS). If out-of-range values are deemed clinically significant, the site investigator(s) must then determine if these values are due to an underlying disease state or constitute an adverse event (potentially even a serious adverse event).

The presence or absence of clinical significance is recorded on the hard copy lab report, which becomes the source documentation. In order to incorporate this information into the clinical database, the reported information can be entered into the database from the lab reports or transcribed onto a CRF and entered with the same processes applied to all other CRFs. Although more

labor-intensive, the latter solution is cleaner and more consistent with other overall study processes. If CRFs are not used, the database should be set up to minimize transcription errors by mirroring the lab report, and may contain some of the following items.

- Local lab name
- Sample collection date (and time, if collected more than once during a visit or for pharmacokinetic analysis)
- Result field for each analyte
- A single "not done" box for the full panel, as well as "not done" boxes for each analyte
- CS/NCS check boxes (evaluation of which is typically the responsibility of clinical reviewers)

Reference ranges (high value, low value and units) and effective dates are collected at the beginning of the study and if reference ranges change. The corresponding lab name should also be collected on a CRF so that during reconciliation between local lab reference range data and the clinical database, as well as for statistical analyses, the results and reference ranges can be merged to create a complete file.

Central Labs

When using a central lab (or any lab that transfers data electronically), the lab and sponsor will complete a data transfer agreement (DTA) during study setup. The DTA defines the format of files, frequency of data transfer, file naming conventions, encryption levels, method of transfer, type of transfer (complete versus partial), recipient, test names, formats, high and low value flags or alerts, and any additional information concerning the lab data. A very important part of the DTA is the definition of data that need to remain blinded. If the result of a certain test could potentially identify which treatment a subject is randomized to or if the subject is responding to treatment, these results need to be blinded. Typically, blinded results remain blank in the file until the clinical database is locked and an unblinding memo is provided. Once this unblinding memo is supplied, the lab releases the information and analysis can occur. The DTA should also include range or data checks being performed by the lab, as well as reconciliation processes.

Cleaning Lab Data

Typical Types of Errors

The most common types of errors from central lab data are demographic errors. When a sample is sent to the lab, a requisition form is completed to identify the subject number, site, sample collection date and time, birth date and gender of the subject (optional) and visit number. If an error is made on the requisition form, this information may differ from the clinical database and prompt a query to be sent to the site or lab.

Careful review and tracking can be used to identify data errors. Review each subject record for values outside defined reference ranges, as well as for consistency of values and units of a given test across multiple visits. If reference ranges are lacking, they should be carefully tracked to ensure all values are associated with the correct reference ranges.[7]

For local lab data not received electronically, the most common errors occur when transcribing results from the printed lab report to the CRF. These errors should be caught by the monitor when reviewing site data, and if caught by the monitor, will not directly impact data management personnel.

Other types of errors encountered may include:

- Interchanged values—Certain values are particularly susceptible to these errors, such as dates, which may be presented differently in the US and Europe.
- Errors in decimal placement—One example would be specific gravity values, which typically have three decimal places (e.g., 1.014). However, sometimes the decimal may be missing, leading to the value being incorrectly recorded as 1014.
- Errors in units—The majority of errors seen in lab data involve inconsistent units. This may happen if different labs are responsible for performing the test for different visits, if the reference ranges and units change during the study, or if the results are recorded in a unit of measurement that differs from that of the reference ranges.
- Misinterpretation of written values, symbols and units—Handwritten numerals, such as 1 and 7, may be misinterpreted due to illegible handwriting on the CRF.

Self-evident Corrections

Self-evident corrections (SECs) are not applicable for electronically transferred data (typically central lab data) but can be used for local lab data if agreed to by all responsible parties. When using local labs, reference ranges should be collected at the beginning of the study for each local lab used at each site. The corresponding lab name should also be recorded on the CRF so that during reconciliation between local lab data and the clinical database, as well as for statistical analyses, the results and reference ranges can be merged to create a complete file. If the lab name on the CRF has been entered incorrectly or misspelled, an SEC can be performed to enter the correct lab name. In order to apply an SEC, the data manager should carefully examine the data to ensure there are no doubts as to the correct information. This will ensure that the correct reference ranges are merged with the corresponding results.

Cleaning Local Lab Data

Lab data recorded on paper CRFs should be subjected to the same data cleaning and edit check specifications as other CRF data, but extra attention should be devoted to verifying subject and lab vendor identifiers. If a local lab transfers data electronically, the measures described in the following section on central lab processes should be adopted.

Cleaning Central Lab Data

Once a test transfer is received, the sponsor or their designee should perform a quality control check of the data

against the DTA to ensure completeness and adherence to the defined structure. If the test transfer is acceptable, regular transfers can begin and reconciliation with the clinical database can commence. The key parameters for reconciliation are information such as subject ID, subject initials, visit or collection date, visit number, visit name, sex, date of birth or age, and test or panel name, although some of these parameters may be optional.

If discrepancies are observed during reconciliation, a query should be sent to the clinical site to verify or correct the information in question. If the query is returned from the site indicating data in the clinical database are correct, the lab data need to be updated according to agreements made with the lab. Some organizations may reverse the order of this process by querying the lab prior to querying the site. When having central lab data corrected or updated, the information should be sent to the lab on a correction log and the lab should update the log once the correction to the data file has been made. This log not only serves as internal documentation during an audit, but also provides the lab with documentation as to why the change was requested and who requested the change. When the changes are made at the lab, a newly updated data file should be sent and reconciliation programs run again. This cycle should occur after every lab data transfer until the data are clean and the clinical database is locked.

Edit Checks for Lab Data

Some standard edit checks that can be applied to lab data include:

- Invalid specimen dates or times
- Blank data, including lab names
- If collecting clinical significance, flagged or out-of-range lab data should be appropriately identified and an associated adverse event should be recorded, when applicable.
- Instances when one test value requires another test value to be provided. For example, if the total bilirubin is greater than 1.0 mg/dL, a direct bilirubin value should be provided.
- Inclusion/exclusion criteria involving lab data can be programmed into edit checks, where appropriate, for flagging when values exceed protocol- defined criteria.
- Listings should be used to compare abnormal results to medical history, adverse events, or other appropriate data.

Lab Accreditation/Certification

According to the International Organization for Standardization (ISO), accreditation is determined as "a procedure by which an authoritative body gives formal recognition that an organization or a person is competent to carry out specific tasks," whereas certification is defined as "a procedure by which a third party gives written assurance that a product, process, or service conforms to specific requirements."[8]

Clinical Laboratory Improvement Amendments (CLIA)

In the US, the term "accreditation" refers both to authorization of labs and to certification of procedures and processes. In 1988, Congress passed CLIA to establish quality standards for lab testing regardless of where the test was performed. The requirements are based on test complexity rather than the type of lab where the testing is performed and are intended to ensure the accuracy, reliability and timeliness of subject test results.

CLIA requires all facilities that perform even one test on "materials derived from the human body for the purpose of providing information for the diagnosis, prevention, or treatment of any disease or impairment of, or the assessment of the health of human beings" to meet certain federal requirements. If a facility performs testing for any of these purposes, it is considered a lab according to CLIA and must obtain a certificate from the CLIA program. CLIA also requires an inspection by the state Department of Health or an accreditation organization such as the College of American Pathologists.[9]

International Accreditation/Certification

The development of quality systems in medical labs of the European Union is based on adherence to the requirements of ISO standards (primarily ISO 15189:2007). The process of accreditation in most European countries is carried out by cooperation among national accreditation bodies, medical experts appointed by scientific associations and health departments. This collaboration has proven successful in the UK, Germany, Hungary, France and Croatia.

Regulatory Agencies

Although it is not a legally binding document, *ICH Guidelines for Good Clinical Practice* provides a solid framework for determining what lab-related documentation should be retained for a study. The regulatory requirements of individual countries will in most cases be very similar to these guidelines, and in some cases the regulatory agencies may be less stringent. Although the ICH guidelines are a great resource, CDM personnel should always consult the regulations of the country in which the study is being conducted. Information regarding regulations from various countries can be found at http://www.hhs.gov/ohrp/international/HSPCompilation.pdf.

For all studies using lab data, *ICH Guidelines for Good Clinical Practice* recommends the following information be kept in the files of the investigator/institution and sponsor.

- Reference values or ranges for all medical/lab/technical procedures or tests
- Changes or updates to reference values or ranges for all medical/lab/technical procedures or tests
- Documentation of certification, accreditation, established quality control, or other validation (where required) of all medical/lab/technical procedures or tests
- Documentation of changes or updates relating to certification, accreditation, established quality control,

or other validation (where required) of all medical/lab/technical procedures or tests[10]

Recommended Standard Operating Procedures

- Data Cleaning
- Laboratory Data Entry
- Laboratory Data Transfers

Competing Interests

The author has no competing interests to declare.

References

1. **Peltonn L, McKusick VA.** Dissecting Human Disease in the Postgenomic Era. *Science*. 2001; 291: 1224–1229. DOI: https://doi.org/10.1126/science.291.5507.1224
2. **Shargel L, Yu A.** *Applied Biopharmaceuics and Pharmacokinetics*. Norwalk, CN: Appleton-Century-Crofts. 1985; 25 ff.
3. **Scherer SW, Cheung J, MacDonald JR, Osborne LR, Nakabayashi K, Herbrick JA,** et al. Human Chromosome 7: DNA Sequence and Biology. *Science*; (April 10, 2003).
4. **Young DS.** Implementation of SI Units for Clinical Laboratory Data. *Annals of Internal Medicine*. 1987; 106: 114–129. DOI: https://doi.org/10.7326/0003-4819-106-1-114
5. **International Conference on Harmonisation.** *ICH Expert Working Group: Guideline for Good Clinical Practice, E6*, Section 8.2.11; 1996.
6. **International Conference on Harmonisation.** *ICH Expert Working Group: Guideline for Good Clinical Practice, E6*, Section 8.3.7; 1996.
7. **Joshi B.** 'Lab Data' in Clinical Trials; Evolving Perspectives. *Data Basics*; 14: 3, pgs. 13–15.
8. **Gligic L.** Status of Development and Implementation of Medical Laboratories Accreditation in Serbia. *JMB*. 2008; 27: 144–147. DOI: https://doi.org/10.2478/v10011-008-0007-2
9. **US Department of Health and Human Services.** *Clinical Laboratory Improvement Amendments (CLIA) Overview.* Available at http://www.cms.hhs.gov/CLIA. Accessed January 8, 2009.
10. **International Conference on Harmonisation.** *ICH Expert Working Group: Guideline for Good Clinical Practice, E6*, 1996.

Further Reading

Eastman JR, Bixler J. Serum Alkaline Phosphatase: Normal Values by Sex and Age. *Clinical Chemistry*. 1977; 23(9): 1769–1770.

Prokscha S. *Practical Guide to Clinical Data Management*. New York, NY: CRC Press. 2007; 75–88.

Snelling LF. Local Lab Ranges. *Data Basics*. 2008; 14: 3, pg. 17.

Chapter Revision History

Publication Date	Comments
October 2009	Initial publication.

How to cite this article: Laboratory Data Handling. *Journal of the Society for Clinical Data Management*. 2023; 1(1): 15, pp. 1–8. DOI: https://doi.org/10.47912/jscdm.335

Submitted: 01 December 2023 **Accepted:** 01 December 2023 **Published:** 22 December 2023

Journal of the Society for Clinical Data Management is a peer-reviewed open access journal published by Society for Clinical Data Management.

OPEN ACCESS