

Unit 5: Quality Assurance, Quality Control and Clinical Data Management

Free notes at pharmacolearn.fyi

5.1 Quality assurance and quality control in clinical trials: the distinction

Quality control, QC, is the operational techniques and activities undertaken within the quality assurance system to verify that the requirements for quality have been fulfilled: the checks performed on the work itself, monitoring in its verification aspect, data checks and edit checks, document completeness checks.

Quality assurance, QA, is the wider, independent system: all the planned and systematic actions, established in advance, that give confidence the trial will be performed and the data generated in compliance with GCP and the applicable requirements, standard operating procedures, training systems, audits, vendor qualification.

Aspect	Quality control	Quality assurance
Nature	Checks on the work as it is done	The planned system that makes good work likely
Performed by	The operational team and its verifiers	A function independent of the operations assured
Typical instruments	Monitoring verification, edit checks, completeness checks	SOPs, training systems, audits, vendor qualification
Timing	In real time, inside the process	Before, during and after, across processes
Failure it prevents	The individual error reaching the record	The same error recurring at every site

5.2 Types of audits

An audit is a systematic and independent examination of trial related activities and documents to determine whether the evaluated activities were conducted, and the data recorded, analysed and accurately reported, according to the protocol, the sponsor's procedures, GCP and the applicable regulatory requirements. The types classify on two axes. By timing and trigger: routine or system audits, planned on a schedule against the audit programme, examining a site, a vendor, a system or the TMF; for cause audits, triggered by a signal, repeated monitoring findings, data anomalies, a complaint; and pre inspection audits, which rehearse an announced regulatory inspection while findings can still be corrected. By object: site audits; sponsor and CRO system audits; vendor audits, of laboratories, data and product suppliers; document and report audits, of the TMF, the database or the clinical study report against their sources; and database audits within data management, the subject of Section 5.15.

5.3 Audit criteria, the audit process and stakeholder responsibilities

Audit criteria are the standards against which the audited activity is judged, fixed before the audit begins. For a clinical trial the set is stable: the protocol and its amendments; ICH GCP; the applicable regulatory requirements, in India the New Drugs and Clinical Trials Rules, 2019 and the related framework; the sponsor's and CRO's standard operating procedures; the trial's own plans; and the agreements governing the work. The audit process runs its recognised course: planning, with scope, objectives, criteria and auditors set in an audit plan and the auditee notified as the plan requires; preparation, studying the trial's documents, monitoring reports and data signals; conduct, with an opening meeting, interviews, facility and

product inspection, and above all tracing of documents and data, consent against source, CRF against source, file against the essential documents list, following each thread until it holds or breaks, findings recorded as objective evidence; and closing, reporting and follow up in Section 5.4. Responsibilities divide accordingly. The sponsor maintains an independent audit programme, appoints qualified auditors, receives reports without interference in their findings, and acts on them. The auditee provides access to premises, documents, systems and staff, answers factually, and owns the corrective work that follows.

5.4 Audit follow up, documentation, resolution and audit certificates

An audit is completed after the visit, in its follow up. The audit report is issued promptly, recording scope, criteria, method, the evidence examined and each finding with its grading and its requirement, and is distributed as the audit procedure prescribes, with the report and its working papers retained in the quality records, the audit documentation, under the confidentiality the material demands. Follow up begins with the auditee's response: for each finding, a root cause analysis that goes past the immediate error to the process that allowed it, and a corrective and preventive action plan, the correction of the instance, the correction of the cause, and the prevention of recurrence elsewhere, with owners and dates proportioned to the grading. Resolution is the state in which every finding stands corrected, prevented or formally accepted at the proper level of management with its risk understood, and the audit is closed. Where the procedures provide for it, an audit certificate, a brief statement that an audit of the named trial was conducted on the stated dates by the sponsor's QA, accompanies the essential documents in the TMF, while the report itself remains within the quality system.

5.5 Preparing for regulatory inspections, including FDA inspections

A regulatory inspection differs from an audit in authority and consequence: it is conducted by the regulator's inspectors, in India the officers acting under the Drugs and Cosmetics framework, in the United States the Food and Drug Administration, whose inspections the syllabus names, and its findings can affect approvals, licences and the acceptability of the trial's data itself. The permanent layer is inspection readiness as a state, not an event: essential documents complete and retrievable, the TMF current, staff who can explain their own work and produce its records, deviations and their handling honestly documented, because inspectors are most persuaded by a system that plainly runs the same way when no inspection is coming. The event layer begins on announcement: a host and presenting roles are named, requested records are produced promptly and logged with copies kept, testimony is checked for consistency with the documents, and staff answer the question asked, from knowledge and documents, without speculation and never with concealment. After the inspection, its report and any observations are answered in writing with corrections and commitments, tracked like the audit follow up of Section 5.4.

5.6 Fraud and misconduct management

Fraud and misconduct management is the quality system's gravest topic, and audits and inspections are where it is most often surfaced. Misconduct covers deliberate wrongdoing and serious negligence in proposing, performing, recording or reporting research: fabrication, inventing data or results; falsification, manipulating materials, processes or records so the research is misrepresented; and the serious breaches that accompany them, consent obtained by deception, eligibility invented, safety events suppressed. Honest error and genuine scientific disagreement are not misconduct. Management follows a disciplined course. Detection comes from the instruments of this subject: source verification that does not reconcile, data and statistical signals, impossible values and improbable uniformity, audit findings, and reports by staff. Assessment is conducted independently and on evidence, the original records secured before they can be altered, the scope established, and participant safety answered first, because participants exposed under fabricated eligibility or concealed risk may need clinical review before any institutional question is

settled. Action is proportionate and documented: notification of senior quality and legal functions, the ethics committee and, as the rules and seriousness require, the regulatory authority; suspension of the site's enrolment or of the implicated work while facts are established; correction of the record and exclusion of unreliable data, with the decision and its basis documented for the analysis and the report.

5.7 Data management: infrastructure and system requirements

Clinical data management, CDM, is the discipline that converts the observations of a trial into an analysable database whose every value can be traced to its source. It begins with infrastructure and system requirements, because data quality is largely decided by the systems chosen before the first datum arrives: the validated systems on which the data live, the access controls, individual accounts, role based permissions and audit trails that make every change attributable, tested backup and recovery arrangements, and the supporting coding, laboratory and safety systems. System requirements are specified before selection: the functions the trial needs, capacity for its size and complexity, compliance with the electronic records requirements applicable to regulated trials, validation status and documentation, vendor support and the vendor's own quality standing. In a regulated trial an unvalidated system is a liability, because data whose system cannot prove who entered, changed or saw them cannot carry the trial's conclusions.

5.8 Electronic data capture systems

Electronic data capture, EDC, is the system family in which trial data are entered directly into electronic case report forms at the site, replacing the paper CRF's journey to a central data entry office with entry at the point of care and checking at the moment of entry. The eCRF screens mirror the CRF design of Section 3.5 and enforce it: fields accept only their defined formats and ranges, edit checks fire at entry, an out of range value, a date out of sequence, a required field empty, so many errors are corrected by the person who made them while the source is in their hand. Queries, the system's questions to the site about unclear, inconsistent or missing data, are raised, answered and closed inside the system with a full audit trail. Access is by individual account and role, investigator, coordinator, monitor, data manager each seeing and doing what their role permits, and the investigator's electronic signature attests the data as the paper signature did.

5.9 Selection and implementation of new systems, validation and test procedures

The selection and implementation of new systems for data management follows the regulated software lifecycle. Selection begins from documented user requirements and proceeds by structured evaluation: candidate systems and vendors assessed against the requirements, the vendor's quality system and support examined, demonstrations tested against realistic trial scenarios, and the total cost, licences, validation, training, migration and support. System validation is the regulated heart of the topic: documented evidence that the system does what it purports to do and continues to do so, built on a validation plan, requirements traced to tests, installation, operational and performance qualification as the organisation's procedures define them, and test procedures that are the executable form of that evidence: written test scripts with expected results, executed on the configured system, outcomes recorded, deviations resolved and retested, the whole retained as the validation documentation.

5.10 Coding dictionaries, data migration and archival

Coding dictionaries are the standard terminologies into which free medical language is converted so data can be counted: adverse events and medical history are coded in MedDRA, the medical dictionary in standard use, whose hierarchy runs from the verbatim term the investigator wrote, through preferred terms, to the system organ classes under which events are tabulated; and medications are coded against the drug

dictionary in standard use, which classifies products by their ingredients and therapeutic classes. Data migration is the movement of trial data between systems, at system replacement, at a CRO transfer, or at the end of a study platform, performed as a validated process in its own right: the mapping of every field and code specified, the migration executed in a test environment and verified, record counts, checksums and sampled value comparisons evidencing that nothing was lost or altered in meaning, and the production migration repeated under the proven procedure. Data archival closes the lifecycle: the trial database and its supporting records, audit trails, coding decisions and validation documents, are archived at study end in a form that preserves readability and integrity for the retention period, with the migration path for ageing formats planned.

5.11 Clinical trial data management: SOPs and the data management plan

The working half of data management is governed by two instruments. Standard operating procedures, SOPs, are the organisation's standing written instructions for every recurring data management activity: study set up, CRF design, database build and validation, data entry and receipt, query management, coding, external data handling and database lock, written, trained and followed so the work does not depend on who performs it. The data management plan, DMP, is the study specific document that applies those SOPs to one trial: written before or alongside database build, approved by the responsible functions, and describing how this study's data will be handled from CRF design to database lock, the systems and versions in use, the data flow, the entry, checking, coding and reconciliation procedures, the external data, laboratory and safety, and how each will be received and reconciled, the query process and its timelines, the quality control steps, and the lock and transfer procedures.

5.12 CRF and database design, study set up and data entry

Study set up is the construction phase of data management, in which the protocol is converted into a working database. CRF and database design considerations start from Section 3.5's derivation: the database is designed behind the CRF, variables defined with formats, ranges and code lists, visits and forms structured as the protocol schedules them, and consistency designed in. The built database is then tested as a system, with test data entered against the specifications, checks fired and verified, and user acceptance testing performed and documented, before release to production and the issue of user accounts by role. Data entry follows the trial's model: in paper trials, entry at the data centre from the received CRF, commonly with independent double entry whose comparison catches keying error; in EDC trials, entry at the site by the trial team, with the system's checks at the point of entry.

5.13 CRF tracking, corrections and data cleaning

Data cleaning is the iterative process that brings the entered data to the state in which analysis is defensible. CRF tracking and corrections, in paper operation, log every CRF page received against the pages expected for every participant and visit, chase the missing, and control corrections: a site corrects its CRF under the correction discipline, original entry preserved and legible, change initialled, dated and explained as required, and the corrected data verified in the database. In EDC operation the same duties live in the system: form status tracking replaces the page log, and corrections are the site's data changes under full audit trail. Data cleaning proper runs on the accumulating database in cycles: programmed edit checks and data listings are run; discrepancies are identified, values failing ranges, dates breaking sequence, results inconsistent across forms; data queries are raised to the sites, each query specific, answerable and tracked; sites respond from source, correcting the data or explaining why the value stands; responses are reviewed, queries closed or reissued; and the cycle repeats as data accrue.

The data cleaning cycle, repeated until lock

Data entered and edit checks run: ranges, dates, cross form consistency, missing items flagged

|

v

Discrepancies reviewed by data management: genuine errors separated from true unusual values

|

v

Queries raised to the site, answered from source documents, data corrected or value explained

|

v

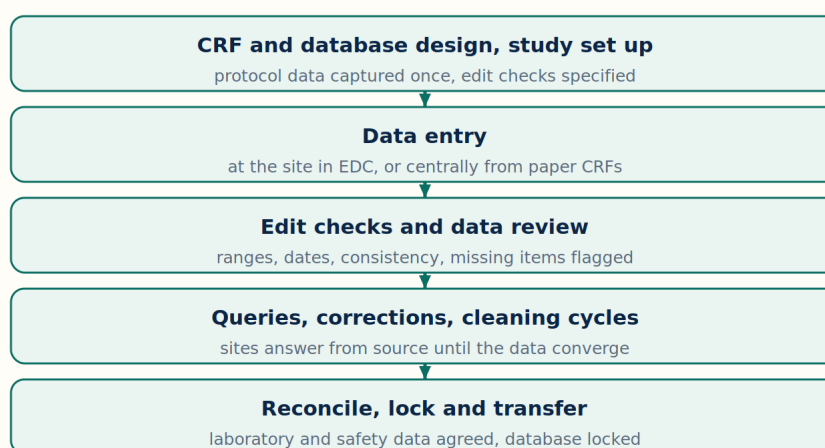
Responses reviewed, queries closed or reissued, coding and reconciliations brought up to date

|

v

Cycle repeats on new data until the database is clean, complete and ready for the lock checklist

Clinical data management: the flow



Cleaning runs with the trial, so lock is a checklist and not a rescue

© 2026 PharmacolLearn, pharmacollearn.fyi

Figure: Data management as a converging cycle. Cleaning runs with the trial, so lock is a short checklist and not a rescue.

5.14 Managing laboratory and ADR data, data transfer and database lock

Laboratory data commonly arrive as external data, transferred electronically from central or local laboratories in agreed formats on an agreed schedule, and are managed by reconciliation: the records received are matched against the participants, visits and samples the trial expected, mismatches queried with the laboratory and the site until the transfer is fully accounted, and the values integrated with their reference ranges and flags so shifts and abnormalities can be reviewed. Adverse drug reaction data live in two systems at once: the clinical database, where events are CRF data coded in MedDRA under Section 5.10, and the sponsor's safety database, where serious events are processed for expedited reporting; the two must agree, and reconciliation between them, event by event for serious cases, is a defined pre lock activity. Data transfer is the controlled movement of the cleaned data to its next custodian, to the statisticians for analysis, between CRO and sponsor, or to an archive, performed under an agreed specification, with transfer verification confirming the data received are the data sent. Database lock is the formal act that ends cleaning: a lock checklist completed, all data received and entered, queries closed, coding finished, reconciliations signed off, protocol deviations reviewed and classified, the analysis populations and any data exclusions decided and documented, access frozen, and the lock authorised in writing by the responsible functions. After lock the database is read only; unlocking, where an error of

consequence is found, is itself a documented, authorised and rare event.

5.15 Quality control and quality assurance in CDM, data mining and warehousing

Quality control in CDM is the checking built into the work: independent review of CRF design and edit check specifications, test evidence for the database build, double entry or its electronic equivalents, query quality review, coding consistency checks, and pre lock QC of the database against the lock checklist, performed inside the data management function as its own verification. Quality assurance in CDM stands outside it, as Section 5.1 defined: the SOP system audited, data management processes and vendors audited, database audits tracing sampled values from source through CRF to database and analysis set, and the metrics of the function, query rates, cycle times, error rates at lock. Data mining is the exploratory analysis of accumulated trial and health data for patterns not specified in advance: safety signals across a programme, subgroup patterns, associations that generate hypotheses; its discipline is the one Section 2.9 taught, that patterns found by searching are hypotheses until a designed study tests them. Data warehousing is the organised, long term repository in which data from many trials are held in common structures, with coding harmonised and provenance preserved, so programme level questions can be answered from data rather than from memory.

Frequently Asked Questions

Q: What is an investigational new drug application and what does it contain? A: It is the application by which a sponsor seeks permission to begin testing an unapproved drug in humans. It assembles the chemistry, manufacturing and controls information, the pharmacology and toxicology evidence with its safety margins, any previous human experience, and the plan for the first studies, the protocol and investigator brochure.

Q: What are the principles of ethics in biomedical research? A: Respect for persons, meaning free and informed agreement with added protection where autonomy is reduced; beneficence with non maleficence, meaning risks identified, minimised, monitored and justified by the expected knowledge; and justice, meaning participants chosen for scientific reasons with burdens and benefits shared fairly.

Q: How is an ethics committee constituted and what does it do? A: It is a multidisciplinary registered body: a chairperson from outside the institution, a member secretary, clinician and basic science members, a legal expert, a social scientist, ethicist or community representative, and a lay person. It reviews the protocol, brochure, consent documents, recruitment and payments, and investigator and site qualifications, approves or requires changes in writing, and continues review of amendments, safety information and progress for the life of the trial.

Q: What are the phases of clinical trials? A: Phase I asks whether the product is tolerable and how the human body handles it, usually in a small number of healthy volunteers, or in patients where the product class requires it. Phase II asks whether it shows activity in patients with the condition, and at what dose. Phase III asks, in larger controlled trials against a comparator, whether it truly benefits patients enough to support approval.

Q: What is the difference between simple randomization, blocking and stratification? A: Simple randomization allocates each participant independently by chance; it is pure but can leave groups uneven in smaller trials. Blocking generates the sequence in small blocks balanced within each block, with random order inside and undisclosed, varied block sizes, so group numbers stay close at every stage. Stratification generates separate sequences within strata of a strong prognostic factor, such as centre or severity, guaranteeing balance within each level of that factor.

Q: How do cohort, case control and cross sectional studies differ? A: A cohort study defines groups by exposure and follows them forward to compare the incidence of the outcome. A case control study defines participants by outcome, cases and controls, and looks back to compare past exposure; it is efficient for rare diseases and long latencies. A cross sectional study measures exposure and outcome at one point in time and yields prevalence and association, but cannot by itself establish which came first.

Q: What documents make up the core clinical trial document set? A: The protocol states what will be done and why. The investigator brochure states what is known about the product, including its risks. The informed consent form records what participation involves, in plain language. The case report form captures exactly the data the protocol promises to analyse. Contracts and agreements bind duties, finances, indemnity and publication rights in writing. Diary cards carry the participant's own record of doses and symptoms between visits.

Q: What happens at a monitoring visit? A: The monitor prepares from the site's data, safety reports and previous findings, then reviews the consent forms for every enrolled participant, verifies reported data against source documents, reviews product storage, accountability and dispensing, checks that procedures followed the protocol, confirms ethics committee communications are current, and looks for safety events recorded in source that were never reported. Findings are recorded in a monitoring report with agreed actions and dates, and followed up until closed.

Q: What is the difference between the Trial Master File, the Investigator Site File and the Pharmacy File? A: The Trial Master File is the sponsor's essential documents for the whole trial across all sites. The Investigator Site File is the site level counterpart held by the investigator, covering that one site's approvals, staff, correspondence, monitoring and conduct records. The Pharmacy File is the trial pharmacy's file for the product at that site: receipts, storage and temperature records, dispensing and accountability logs, and return and destruction documentation.

Q: What is database lock in clinical data management? A: Database lock is the formal act that ends data cleaning and freezes the trial database for analysis. It happens only when a lock checklist is complete: all data received and entered, queries closed, coding finished, laboratory and safety data reconciled, deviations reviewed and classified, and analysis populations and exclusions decided and documented. Access is then frozen to read only and the lock is authorised in writing.

*More free M.Pharm notes and unit wise PDFs at pharmacolearn.fyi
Free video lectures on the PharmacoLearn YouTube channel*