



Module 3.1 Transcript

Interpretation and Application of ICH E6(R3):

GOOD CLINICAL PRACTICE (GCP)

Welcome to the Interpretation and Application of ICH E6(R3): Good Clinical Practice.

For instructions on how to navigate the system, please select the type of device you're using to view this module.

Click next to begin.

Module 3 of this training discusses data governance in clinical trials.

This module contains 4 sub-modules, which should be taken sequentially. We'll begin with Module 3.1: Introduction to Data Governance.

Key questions you will be able to answer after completing this module include:

- What is data governance and why is it important in clinical trials?
- Why was a new section on data governance introduced in ICH E6(R3)?
- And,
- How does the data governance section align with sponsor and investigator responsibilities in ICH E6(R3)?

Before we begin, let's explain what we mean when we talk about data governance in clinical trials.

Data governance encompasses how data are managed, protected, and maintained throughout the data lifecycle.

Data governance refers to the systematic management of data to ensure its integrity and traceability to its original source.

Data integrity is a new glossary term in ICH E6(R3) and includes the degree to which data fulfil key criteria of being attributable, legible, contemporaneous, original, accurate, complete, secure, and reliable, such that data are fit for purpose.

So, why is data governance important for clinical trial quality?

Principle 9.1 states that the quality and amount of information generated in a clinical trial should be fit for purpose and sufficient to provide confidence in the trial's results and support good decision making. Effective, risk-proportionate data governance ensures this principle.

In other words, data governance lays the foundation for generating trustworthy and reliable clinical trial results. Together with quality processes and systems, effective, risk-proportionate data governance is essential to achieving clinical trial quality that is fit for purpose.

Reliable and trustworthy clinical trial results are the foundation for sound clinical decision making.

The E6(R3) guideline applies to interventional trials of investigational products that are intended to be submitted to regulatory authorities. The guideline also explains that the principles of GCP may also apply to other interventional trials of investigational products that are not intended to support marketing authorisation applications. Whether GCP principles apply to those trials not intended to support marketing authorisation applications depends on local regulatory requirements.

Even though the guideline is mainly directed toward trials intended to be submitted to regulatory authorities, the expectations for data governance described in E6(R3) are important to consider in any interventional trial for generating reliable results.

These expectations support regulatory decision making and help inform clinical decision making.

ICH E6(R3) introduces, for the first time, a dedicated section on data governance. It was added for four important reasons:

- First, a stronger focus on data governance is intended to strengthen clinical trial data quality, integrity, and reliability.
- Second, this new section supports clinical trial modernisation by providing a framework that enables the adoption of innovative technologies in clinical trial design and conduct.
- Third, the new section, together with the relevant sections on the investigator and the sponsor, clarifies roles and responsibilities, defining expectations for sponsors and investigators in managing data, systems, and processes throughout the data and computerised system lifecycles to ensure their quality, integrity, and reliability.
- Finally, the new data governance section is intended to minimise the risks of misinterpreting E6(R3) requirements by providing clear guidance and flexibility for users.

Let's briefly look at the concept of responsible parties and how that relates to data governance responsibilities.

The sponsor and investigator are the two responsible parties in a clinical trial. As such, they are expected to collaborate closely to ensure the integrity, protection, and usability of trial data.

Therefore, Section 4: Data Governance should be read in conjunction with Section 2.12 (Investigator Responsibilities: Records) and 3.16 (Sponsor Responsibilities: Data

and Records). These sections, together with Section 4, clarify the responsibilities and expectations of each party in ensuring GCP compliance related to the quality, integrity, and reliability of trial data and records.

Reflecting Principle 10, sponsors may transfer and investigators may delegate activities related to data governance to service providers.

“Service provider” is newly defined in the ICH E6(R3) glossary as “a person or organisation (commercial, academic or other) providing a service used by either the sponsor or the investigator to fulfil trial-related activities.”

Tasks performed by service providers related to data, computerised systems, and data governance processes might include, among others, building, hosting, or otherwise managing computerised systems or trial data.

Another new term in the ICH E6(R3) glossary is “Data acquisition tool” or “DAT” for short.

A data acquisition tool is defined as “a paper or electronic tool designed to collect data and associated metadata from a data originator in a clinical trial according to the protocol and to report the data to the sponsor. The data originator may be a human (e.g., the participant or trial staff), a machine (e.g., wearables and sensors) or a computer system from which the electronic transfer of data from one system to another has been undertaken (e.g., extraction of data from an electronic health record or laboratory system).”

Examples of data acquisition tools include:

- Case report forms (CRFs)
- Clinical outcome assessments, including patient-reported outcomes
- And wearable devices.

We will explore in later modules the complementary and sometimes overlapping responsibilities that the investigator and sponsor have for the management of data throughout the data lifecycle, from data capture, through data transfer, all the way to destruction.

We will also explore the management of computerised systems throughout the computerised system lifecycle.

We will look at these lifecycles and relevant responsibilities of the investigator and sponsor in Modules 3.3 and 3.4.

Throughout the rest of Module 3, we will use an in-depth case example to explore how a risk-proportionate approach can be applied to data governance activities and to highlight key data governance responsibilities for sponsors and investigators. This case example will be introduced here and explored in greater detail in the upcoming modules. Let's learn about the case example.

The trial we will be using as our case example is a multicentre, randomised, double-blind, placebo-controlled induction trial to assess the efficacy and safety of an investigational product in participants with moderately to severely active Crohn's Disease who failed prior biologic treatment.

The primary trial objective is to evaluate the efficacy and safety of the investigational product versus placebo during induction therapy in participants with moderately to severely active Crohn's Disease who have failed a prior biologic treatment.

The trial has two co-primary efficacy endpoints, meaning that both endpoints must be met for the trial to be successful.

The first is Crohn's Disease Activity Index Clinical Remission (or CDAI clinical remission) at Week 12. This is measured by the proportion of participants achieving a CDAI score of less than 150. The second co-primary endpoint is endoscopic response at week 12.

Assessment and data collection for the CDAI co-primary endpoint is made up of 8 components that are combined using a standardised formula in which each component is assigned a weight reflecting its relative contribution to disease activity. The weighted components are summed to generate the total CDAI score.

The eight components fall into 3 groups.

- First, participants report symptoms daily using an electronic patient-reported outcome tool, or E-pro and each day record the number of liquid or soft stools, their level of abdominal pain, and their general wellbeing. The CDAI score is subsequently calculated using eDiary data from 7 days prior to each visit.
- Second, a clinician collects additional information at baseline and every 4 weeks thereafter. The clinician records whether there are any extraintestinal complications, whether anti-diarrhoeal medications are being used, whether an abdominal mass is present, and whether there has been any weight loss.
- Finally, the participants' haematocrit levels are measured at baseline and every 4 weeks thereafter.

The second co-primary efficacy endpoint is Endoscopic Response at Week 12, defined as a greater than 50% decrease in Simple Endoscopic Score for Crohn's Disease from baseline.

Endoscopic response is assessed at baseline and Week 12 using an endoscopy performed at the investigator site, with all evaluations of the images performed by a blinded central reviewer.

The co-primary efficacy endpoints involve multiple complex data flows, using different data acquisition tools and other computerised systems, data collection processes and service providers, each carrying distinct risks that require tailored oversight and controls.

Such trial complexity warrants risk-proportionate data governance across the data lifecycle, including system management, data collection, review, finalisation, and analysis. Because errors in system design, use, oversight, or data transfer could compromise data integrity, these data, systems and processes necessitate more careful review.

In the next section of this module, you'll explore the data flow diagram for each component of the co-primary efficacy endpoint data.

We will break down these data flows and explore the data, systems, and processes that are of higher criticality.

Click the green circle to continue.

Click on "Patient-reported CDAI components" to begin.

Participants record their symptoms each day using an electronic diary, or eDiary. They report stool frequency, abdominal pain, and general well-being.

These entries make up the patient-reported components of the CDAI. This information is collected as an electronic patient-reported outcome, or ePRO, which is a type of electronic clinical outcome assessment, or eCOA.

The eCOA device automatically transfers data after the completion of a diary entry to the eCOA Service Provider Database.

From there, the protocol requires delegated investigator site staff to review the ePRO data captured in the eCOA database and transcribe daily values for the 7 days prior to each visit into the eCRFs at baseline and every 4 weeks thereafter. During

transcription, investigator site staff must exclude days associated with bowel preparation, endoscopy procedures, and the two days following endoscopy.

This transcription process adds operational complexity and increases the risk of data handling errors.

The eCRF is pre-programmed to calculate the CDAI scores from those data and data from the other CDAI components.

Following completion of monitoring and data finalisation activities, eCRF data and derived CDAI scores are compiled and used to generate the study data tabulation model dataset (or SDTM dataset, for short), which is then transferred to the sponsor database.

Click on “Clinician-assessed CDAI components” to continue.

The investigator or delegated investigator site staff collect data on the presence of extraintestinal complications, participant use of anti-diarrheal drugs, the presence of abdominal mass, and weight loss. Data for the clinician-assessed CDAI components are collected at baseline and every 4 weeks thereafter.

The data for these components are entered directly into the eCRF by the investigator or investigator site staff.

In the same manner as we saw in the patient-reported data flow, the eCRF is programmed to calculate the CDAI score from the data entered.

As before, the eCRF data are compiled and used to generate the SDTM dataset, which is then transferred to the sponsor database.

Click on “Laboratory CDAI component” to continue.

Haematocrit values are the final component of the CDAI score. The data flow for this component begins when a blood sample is taken by delegated investigator site staff at baseline and every 4 weeks thereafter.

The blood sample is sent to a central laboratory for processing and analysis. Following analysis, the results are automatically transferred from the laboratory analyser to the laboratory information management system. The lab results are then reported back to investigator site staff via a report provided by the central lab.

Investigator site staff then enter the haematocrit values into the eCRF.

As with all other CDAI components, the eCRF uses this data to calculate the CDAI score.

As before, then data from the eCRFs are used to generate the SDTM dataset, which is then transferred to the sponsor database.

Click on “Endoscopy data” to view the final data flow diagram.

Now, let’s look at the data flow for the second co-primary endpoint, endoscopic response. The endoscopy procedure is conducted by the investigator or a delegated gastroenterologist at baseline and Week 12.

Investigator site staff then upload the endoscopy images to the Central Imaging System.

Next, images are assigned to central readers in a blinded fashion. The central readers then evaluate the images and enter scoring into the Central Imaging System.

Following completion of monitoring and data finalisation activities, the imaging-related data are transferred to the sponsor for incorporation into the SDTM datasets and for analysis.

That completes the main data flows for this case example.

Thank you for completing Module 3.1.

In the sub-modules that follow, we will use these data flow diagrams to explore the data, systems, and processes of higher criticality as well as their associated risks, potential risk mitigation strategies, and the specific sponsor and investigator responsibilities needed to ensure data quality, integrity, and reliability throughout the data lifecycle.

Next, in Module 3.2, we’ll look at how to take a risk-proportionate approach to data governance.