Data collection in clinical trials

Introduction

When a clinical trial is being designed, it is important to plan how data will be collected and captured during the trial.

This article describes the process of documenting a clinical trial, including:

- Where the data are recorded by the investigator
- How the data are collected
- How all the documents that are generated for a study are compiled for potential inspection by the competent authorities at the sites of the investigator and of the sponsor.

Types of data collection in clinical trials

Data in a clinical trial are generated and collected by:

- The investigator
- Study staff
- Directly by patients (called Patient-Reported Outcomes (PROs))

This can occur in the traditional way – on paper (such as Case Report Forms (CRFs), patient diaries, or questionnaires); or in electronic ways – for instance in electronic CRFs (eCRFs), or by using hand-held instruments such as mobile phones or tablets to collect data directly from patients (ePROs). Another method of collecting data is called ‘direct data capture’ (DDC). In DDC, data are directly generated by

electronic devices and entered into the database

Paper Case Report Forms (CRFs)

Paper CRFs are designed for handwritten data. They are cheap to produce and allow the creation of direct copies and faxing. New technology such as optical character recognition (OCR) allows computers to ‘read’ the data written by site staff and enter them automatically into a database.

Advantages:

- Site staff can carry the CRF to wherever they need it
- Site staff don’t need to worry about access to computers and passwords.
- Relatively easy to amend if changes are required during the study

Disadvantages:

- A large volume of paper to store
- Space and correction limitations on the form itself
- Incorrect data entries are not automatically flagged to the user as they may be on electronic records
- As data is later entered into a database, creates another opportunity for mistakes to be made

Electronic Case Report Forms (eCRFs)

Electronic CRFs (eCRFS) are becoming more and more popular. However, they are much more complicated to produce and need to adhere to strict regulations in Europe and the United States. The computer programmes or software must be validated, and every correction that is made to the data entered must be traceable. They must ensure that only authorised persons have access to the programme and to the data. Data backups must occur regularly and automatically.

Using eCRFs in a study requires all investigator sites to have

sufficient and reliable access to computers and the internet. It also requires intensive training of the site staff using the eCRF, which must often also be supported by a help-desk.

Regulatory requirements are in place that eCRFs must conform to:

- In Europe: ICH GCP E-6, Section 5.5.3[1]
- In the US: FDA – 21CFR Part 11 and Guidance for Industry – Computerised Systems used in Clinical Trials[2]

System validation

The validation of electronic systems is mandatory. A system must:

- Have an audit trail, meaning that any change should be electronically recorded and traceable;
- Be protected against unauthorised access;
- Be backed up regularly, meaning that data are regularly copied on a different disk, server, or computer that can be accessed for the lifetime of the product.

The US Food and Drug Administration has worked out very detailed and demanding rules outlining the conditions under which they accept electronic data capture.

Guidance for industry

Recommends that the protocol should identify when a computerised system will be used to create, modify, maintain, archive, retrieve or transmit data.

Documentation of all software and hardware used should be kept with study records.

Advantages

- Data entry errors are directly detected
- Range and edit checks minimise data entry errors and

protocol violations
- Data are available to sponsor immediately after entry at site
- Faster query resolution is possible

Disadvantages

- Benefit only seen in the long term
- Data entry done by site personnel
- Residual resistance exists to electronic data capture
- Technical problems may occur
- Data protection issues may arise

Examples of Direct Data Capture (DDC)

- Laboratory data
- Electrocardiogram (ECG) data
- Central image reading (Magnetic Resonance Imaging (MRI) results)
- Electronic patient questionnaires / diaries

Patient Reported Outcomes (PROs) and Electronic Captured PROs (ePROs)

The term Patient Reported Outcome (PRO) is used for all data that are directly provided by patients. This includes all types of questionnaires and diaries. This can be recorded on paper or by using electronic systems. Technical tools that can be used to receive these data in an efficient, participant-friendly manner are rapidly evolving. If an electronic hand-held system such as a tablet or text messaging (SMS) is used, the term ePRO is used. Typically, these electronic data are either in the form of a daily diary at the patient’s home, or Quality of Life (QoL) questionnaires administered during site visits.

Advantages

Asking patients to provide their data electronically has many advantages: the quality of data is better, and these systems allow the site staff ongoing understanding of how the patient is doing, and whether the data are entered reliably or not. With paper diaries, this only becomes obvious at the next patient visit when they bring the diary to the site. ePROs also reduces the study data entry workload for the site staff.

Higher quality of data:

- Automated edit checks ensure PRO data is often 100% clean – meaning there is no need for extensive data cleaning
- Alarms and context-sensitive eDiary design achieves much higher compliance to protocol
- Higher quality data might mean fewer patients are needed in a study
- Immediate intervention possible when problems or deviations occur
- Allows clinicians to concentrate on treating their patients rather than on data entry

Disadvantages

There are also a number of disadvantages to consider when including ePROs into a clinical trial. Statistics show that the benefits are more and more dominant because the number of studies involving ePROs is increasing rapidly.

- Higher technical effort and therefore more expensive than paper
- Not all patients are familiar with modern technology
- As with every electronic instrument, there can be failures and break-downs
- More time required for the site staff to explain the use of the system to the patient
- Telephone lines or wireless networks need to be

available

Patient involvement

- PROs give sponsors a structure with which to seek real-life experiences from participants during a trial.
- QoL evaluations include measures of a participant's ability to conduct everyday tasks (for instance, those that they might otherwise find difficult) can provide important findings associated with a participant's experience during a trial. These real-life data often become important in decision-making when a product gets a marketing authorisation and is being assessed by Health Technology Assessment (HTA) bodies.
- Patient experts (patient organisations or representatives) should therefore be involved in order to define the QoL or other patient data that should be collected. This provides an opportunity for patients to have a greater role.

Conclusions: The importance of high-quality data

Ultimately, however the data in a clinical trial are captured and handled, they must be of the best possible quality. The criteria for high quality data are that they:

- Can be evaluated and analysed
- Allow valid conclusions to be drawn
- Are complete and accurate
- Do not need to be queried
- Are consistent across subjects and sites
- Are complete for all CRF fields
- Are legible and easy to understand
- Make logical sense
- Are in the correct units
- Provide greater clarity around subjective experiences

[glossary_exclude]Further Resources

1. The European Medicines Agency has issued a reflection paper summarising what Good Clinical Practice (GCP) inspectors will accept as electronic data capture: European Medicines Agency (2023). *EMA/INS/GCP/112288/2023 Guideline on computerised systems and electronic data in clinical trials*. Retrieved 18 February, 2024, from https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-computerised-systems-and-electronic-data-clinical-trials_en.pdf
2. S. Food and Drug Administration (2003). *Guidance for industry: Part 11, Electronic records; electronic signatures – scope and application*. Retrieved 7 September, 2015, from http://www.fda.gov/RegulatoryInformation/Guidances/ucm125067.htm
3. S. Food and Drug Administration (2009). *Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labelling claims.* Retrieved 7 September, 2015, from http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm193282.pdf
4. The Comet Initiative comprises researchers interested in the development and application of agreed standardised sets of outcomes: a 'core outcome set'. For more information, see http://www.comet-initiative.org/about/overview and http://www.comet-initiative.org/resources/PlainLanguageSummary[/glossary_exclude]

[glossary_exclude]References

1. International Conference on Harmonisation (1996). 'Trial management, data handling, and record keeping.'

Guideline for Good Clinical Practice E6(R2) (pp. 23). Geneva: ICH. Retrieved 5 July, 2021 from https://database.ich.org/sites/default/files/E6_R2_Addendum.pdf
2. S. Food and Drug Administration (2003). *Guidance for industry: Part 11, Electronic records; electronic signatures — scope and application*. Retrieved 7 September, 2015, from http://www.fda.gov/RegulatoryInformation/Guidances/ucm125067.htm[/glossary_exclude]

[glossary_exclude]Attachments

- Presentation: Options for Data Collection and Patient-reported Outcomes (PROs)
  Size: 611,418 bytes, Format: .pptx
  A presentation covering the various types of data collection in clinical trials.

[/glossary_exclude]
A2-4.28-v1.1