EUROPEAN CLINICAL RESEARCH INFRASTRUCTURES NETWORK - TRANSNATIONAL WORKING GROUPS

ECRIN-TWG

FP6-2005-Life Sciences and Health LSH-2005-3-4
Contract # 037199

Deliverable D10

GCP-compliant data management in multinational clinical trials

Version: **1, final**

Date of preparation: **15 September 2008**

Work package 4: **Transnational Working Group on Data Management**

Work package leader: Christian Ohmann
Coordination Centre for Clinical Trials
Medical Faculty
Heinrich-Heine-University
Moorenstr. 5
40225 Düsseldorf
Germany
+49-(0)211-81-19701
+49-(0)211-81-19702
email: Christian.Ohmann@uni-duesseldorf.de
http://www.uniklinik-duesseldorf.de/kks

Working group members and authors:

Dimitrios Venizeleas, Wolfgang Kuchinke, Coordination Centre for Clinical Trials, Heinrich- Heine- University, Düsseldorf, Germany

Nader Salas, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Jens Lauritsen, Orthopedic Department, Odense University, Hospital and Institute of Public Health, Dept. of biostatistics, Southern Denmark University. Odense, Denmark

Francois Gueyffier, Centre d'Investigation Clinique, Hôpital Cardiologique Louis Pradel, Lyon Cedex 03, France

Alan Leizorovicz, Faculté RTH Laennec, Service de Pharmacologie Clinique, Lyon Cedex 8, France

Carmen Schade-Brittinger, Maik Hahmann, Coordination Centre for Clinical Trials, Philipps- Universität Marburg, Germany, supported by data managers of the KKS-network, Germany

Zoltán Voko, Division of Epidemiology and Biostatistics, Department of Preventive Medicine, Debrecen University, Hungary

Siobhan Gaynor, Margaret Cooney, Molecular Medicine Ireland

Peter Doran, University College Dublin, Clinical Research Centre, Mater-Misericordiae University Hospital and St. Vincent's University Hospital, Dublin, Ireland

Aldo Maggioni, Andrea Lorimer, Centro Studi ANMCO, Firenze, Italy

Ferràn Torres, Hosptital Clinic i Provincial, Barcelona, Spain

Joseph M Castel, Hospital Universitari de la Vall d'Hebrón, Barcelona, Spain

Gladys McPherson, Health Services Research Unit,

University of Aberdeen, UK

Jim Charwill, UK Clinical
Research Network Coordinating Centre, Leeds, UK

Stuart W. Cummings, PSI Pharma support International,
Brussels, Belgium

Mats Hellström, Department of Oncology, Karolinska
University Hospital Sweden

Stéphane Lejeune, European Organisation for Research and
Treatment of Cancer (EORTC), Bruxelles, Belgium

CONTENTS

1. Introduction

Compliance with 'Good Clinical Practice (GCP)' is a prerequisite for many clinical trials (e.g. trials with medicinal products) and it is increasingly recommended for all types of trials. GCP is an international, ethical and scientific quality standard for designing, conducting, recording and reporting trials that involve the participation of human participants. Compliance with this standard provides public assurance that the rights, safety and well-being of trial participants are protected and that it is consistent with the principles that have their origin in the Declaration of Helsinki. Moreover, compliance with GCP is key to ensuring the credibility of trial findings.

The implications of application of this standard on data management have to be defined in the European context, where different national regulations must be integrated into principles and requirements that apply Europe-wide. In this context it is a key role of the European Clinical Research Infrastructures Network (ECRIN) to provide leadership in the application of the various European standards and the articulation of a GCP-compliant data management policy. Such a policy, upon adoption by all ECRIN members, will become a roadmap to ensure the credibility of ECRIN-conducted multinational trials.

The term "GCP compliant data management" has to be clarified as well in the context of multinational multicentre clinical trials. Credible data means that data have to be trustworthy, i.e., they have their origin in observations of study participants as reported in source documents and are not altered or falsified. Data have to be complete, accurate and verifiable. Since there is no definition of "GCP-compliance" for data management, we have first evaluated the necessary regulations and guidelines to find the relevant sections in them (see Appendix II). From these documents, recommendations will be defined regarding all areas of clinical trial data management, with the ultimate objective of defining the processes, software tools, platforms, interfaces and development models that comply with these standards in the ECRIN environment.

2. Objectives and scope

The objectives of this document are to present recommendations within a framework for data quality assurance of clinical trials and to develop a foundation for harmonised interpretation of GCP requirements for data management within the ECRIN framework. In addition, we will seek to identify good data management practices in general. Core aspects are the identification of relevant regulatory information, industry standards and good practice in all aspects of data management in regulated clinical trials.

The major areas to be addressed within this document include:

- Consensus on definition of GCP as it applies in the data management environment
- Dissemination of information regarding regulatory standards and requirements to all ECRIN members
- Definition of the processes that need to be developed to ensure compliance with GCP

- Creation of a set of recommendations for quality data management
- Description of a framework and an approach for moving on in the discussion of data management issues within ECRIN, resulting in deliverables that can be used for evaluation and audit of data management centres applying for participation in ECRIN-based trials
- Specification of what is required to meet the requirements of ECRIN for supplying data management services and providing comprehensive and detailed information to centres interested in becoming an ECRIN data centre
- Assessment of issues including risk analysis and provision of a kind of comprehensive easy-to-use checklist for the evaluation of procedures or units.

Quality management processes and structures will have to be addressed, as well as the critical issue of system validation. To support the harmonisation process within ECRIN, not only minimal criteria for compliance with regulatory requirements, but also "best practices" should be developed, reflecting a thorough understanding of GCP and its requirements, allowing implementing units to fully comply with GCP-standards in their data management activities.

In this document the minimal requirements vs. best practices will be additionally indicated with the following bullet points:

✓ **Minimal requirement (example)**
☺ **Best practice (example)**

An important question is how to deal with different technical approaches in conducting clinical trial data management, referring mainly to the dualism of electronic versus paper case-report forms (CRF). For this purpose, four different scenarios for data collection were selected which form the basis for technical discussion (see Appendix III: scenarios). This document will be based on the assumption of a general document *"Case Report Form" (CRF)* and only perform a case discrimination (*electronic CRF v. paper CRF, eCRF/pCRF*) when necessary. For example, a Standard Operating Procedure (SOP) regulating clinical data processing describes the process independent of the eCRF/pCRF question and provides two 'sub-sections' '*Electronic CRFs, Paper CRFs'.* The objective of this document is not to give technical recommendations of how to implement electronic trials, but rather to provide GCP requirements needed to specify implementation guidelines for the trial, be it paper based or electronically performed.

3. Identification of core requirements

To identify the core requirements for a data management system, a traceability matrix approach is used. Each aspect of the system is referenced to the corresponding regulatory source.

There are different regulatory documents and guidelines from which requirements can be derived. The most important ones are local, national or European legislation and guidelines. In addition, actual requirements and standards for state of the art technology need to be specified. Recommendations are given from "minimum

requirements" to "best practice". The appropriate level for a given study depends on various factors, such as economy, internationalisation, safety and adverse effects of alternative or standard treatment, logistical possibilities, e.g. working in less affluent areas etc.

The status quo must be adapted to open and documented industrial norms, e.g. those of national or international standardisation committees. Based on these standards and norms, systems can be consequentially developed or adopted, and developers of industrial standards or platforms that have become widely accepted must be influenced to follow the open industrial norms. Of particular difficulty is the issue of long-term archiving for the period demanded of access to the primary data; usually at least 10, but could be up to 15 years. Archiving demands usage of data, independent of specific IT-systems or database formats.

Sometimes data management practice cannot be based on existing regulation because many areas are not regulated explicitly. But the regulations have underlying concepts that recommended practices should aim to be consistent with.

The most important sources for GCP-compliant data management referring to the EU are the following:

- ICH GCP
- EU Directive 2001/20/EC
- EU Directive 2005/28/EC.

Additional sources refer to regulations in the US or to other recommendations, such as

- FDA Computerized System Used in Clinical Trials
- Good Clinical Data Management Practices, version 4
- CDISC Clinical Data Interchange Standards Consortium, Operational Data model (ODM)
- CONSORT documentation for reporting of clinical trials.

In general, regulatory requirements rarely directly refer to data management processes. It is, therefore, the scope of this document to extract the necessary information from the different relevant sources into a consistent framework of rules and best practice recommendations. Obviously, there is place for discussion and for having different opinions; therefore, the framework should be stringent enough to find a common, solid basis for conducting interoperable data management within the ECRIN network. On the other hand, the framework should provide sufficient flexibility to be able to deal with different requirements, e.g. due to local legislation. This consensus can be built through recognition of the fact that best quality approaches should be the primary focus.

4. Specific FDA issues

The FDA is the US Government regulatory office for registration of pharmaceutical products. Here especially the Code of Federal Regulations (CFR) applies, which is the codification of the general and permanent rules published in the Federal Register by the agencies of the Federal Government. FDA regulation is relevant for EU projects in development of drugs considered for possible registration in the US.

Therefore, more specific regulation is available, e.g. for electronic documentation, the consideration of the US regulation is particularly helpful. However, it must be clarified, that in the EU it is not the FDA regulations which are governing, but the national implementations of EU directives or the EMEA implementations of EU regulations.

The FDA is primarily concerned about the following aspects of clinical trial data: attributable, legible, contemporaneous, original and accurate. The data reviewed by FDA have to be the original data collected at the investigator's site. The main requirement, therefore, is a **robust audit trail**. The FDA guidance for industry "Computerized Systems Used In Clinical Trials (CSUCT)" (1999) and the Electronic records/Electronic signature rule (21 CFR Part 11) including guidance are important in this regard. In this context, it is important to note that the local implementation of the rules has to be taken into consideration, particularly where national regulations might conflict with FDA requirements (e.g. electronic signature as implemented in today's EDC systems might be inappropriate for German legislation).

The FDA is encouraging the use of computerized systems, but such systems have to meet certain requirements.

The key requirements include:

- system validation is crucial
  - to ensure authenticity, integrity, confidentiality and non-repudiation of data and signed records
  - to ensure accuracy, reliability, consistent intended performance, and the ability to discern invalid or altered records

- system validation has important components
  - requirement specifications
  - validation plan
  - test plan
  - traceability matrices

- requirement specifications are needed for
  - system design
  - edit checks
  - archiving procedures
  - audit trail design
  - security access controls
  - authenticity controls
  - privacy controls

- the agency accepts three ways of signing
  - digital signature
  - biometric signature
  - handwritten signature
- if data are collected electronically, they should be archived electronically

- the archived data should enable a reconstruction and evaluation of the trial.

5. Important components to ensure GCP-compliant data management

The following issues have to be considered for the implementation of GCP-compliant data management systems and procedures. But first, some important concepts have to be agreed upon and well understood.

5.1 Basic terms and concepts

5.1.1 General definition of source data and related terms

All the following terms are as defined by the CDISC Clinical Research Glossary (Version 6.0, Applied Clinical Trials, Dec. 2007):

Source data:
All information in original records and certified copies of original records of clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies).

Source documents:
original documents, data, and records (e.g., hospital records, clinical and office charts, laboratory notes, memoranda, subjects' diaries or evaluation checklists, pharmacy dispensing records, recorded data from automated instruments, copies or transcriptions certified after verification as being accurate copies, microfiches, photographic negatives, microfilm or magnetic media, x-rays, subject files, and records kept at the pharmacy, at the laboratories and at medicotechnical departments involved in the clinical trial).

CDISC Clinical Research Glossary, Version 6

eSource data (electronic source data):
Source data captured initially into a permanent electronic record used for the reconstruction and evaluation of a trial. NOTE: "Permanent" in the context of these definitions implies that any changes made to the electronic data are recorded via an audit trail.

CDISC Clinical Research Glossary, Version 6

Thus, eSource data describe source data that are acquired/measured for the first time and then captured initially into a permanent electronic record. If data are entered from another system, in which data were acquired/measured for the first time, it is transferral of eSource data. An example is laboratory data primarily collected electronically as eSource and then transferred to a trial database.

eSource document:
The electronic record used to keep together a collection of eSource data items for capture, transmission, storage, and/or display; and serving as a *source document* for a clinical investigation. NOTE: Electronic source documents are recorded in electronic systems according to conventions (such as those for PDF documents) that ensure that all the fields of eSource data and associated contextual information (e.g., time of capture, time zone, authorship, signatures, revisions) are linked to each other in a certain order of presentation. The encoded specifications in the electronic record thus serve the same role as the physical properties of paper (binding items together). eSource documents are subject to regulations and guidance that apply to source documents.

CDISC Clinical Research Glossary, Version 6

It is recommended that a digital signature mechanism be incorporated into the data capturing to legitimate the registered data as an authorised source data. In this way the data has its own timestamp and responsible person's information.

5.1.2 "eTrial" file

When using software support for the processes of clinical trials the Life Cycle Model should be considered. This model covers the phases of planning, specification, design, construction, testing, installation, acceptance testing and operation. In a wider context, a clinical trial may be seen as part of a research process, covering e.g. background research, grant applications, publishing and translation of research findings. As an example, the National Institute for Health Research (NIHR) Longitudinal Research Record holds all the information about a research study necessary to manage it throughout its full lifecycle. The "eTrial" file, however, addresses only the data management processes, which will apply across several stages of the typical trial lifecycle.

Certain types of clinical computer systems require special attention during maintenance and validation. Such systems include administration, randomization, drug supplies, data capture and transfer, databases and associated review tools, statistical analysis, document management, publishing and regulatory submission and pharmacovigilance. Assuming that at any time of an ongoing or finished trial, an "eTrial"-file can be generated for submission to authorities, backup/archiving purposes, or for statistical analyses, a minimum data standard should be defined for the structure of an "eTrial"- file. For full implementation of best practice some standards can be defined or adopted, e.g. ODM from CDISC.

The mechanism that ensures data integrity and consistency of this file (guarantees that the data have not been changed accidentally or intentionally from the source medium to the destination medium) exists as a digital signature (e.g. on XML files).

When the minimal requirements are fulfilled, a SOP for reformatting of any file type into a standardized format should be devised.

5.1.3 Requirements for acquiring/capturing/copying source data

In general, documents containing source data must first be specified in the trial protocol. Source data are the original data, the recordings and all information regarding clinical investigations, documented for a patient in the clinical record *according to the protocol*. These may be data that resulted from investigations, laboratory findings, anamnesis, interviews, patient diaries and other sources. The original documents have to be archived. Copies have to be dated and signed by a responsible person (certified copies). If the original data are stored electronically, a printout has to be made or a list of dates and versions of stored documents signed/dated by the investigator. In the case of eSource data, of course, this is not possible. A copy of eSource data shall be accepted in place of eSource data, if the copy has been produced and verified against the eSource data based on procedures defined in a SOP for acquiring data duplication and verification.
Appropriate handling is also required for scanning source documents. The scanning process has to be validated prior to implementation in a trial to ensure the integrity of the generated record.
If the CRF is the source document (e.g. in psychiatric instruments like psychometric scales) this has to be defined in the protocol. If work sheets have been used as a transcription instrument (e.g. transitional documentation prior to electronic data entry), these are to be considered as informal source data sheets and have to be filed and quality checked appropriately.
In general, source data must be accessible and verifiable and the quality of digitisation must be carefully controlled using appropriately defined SOPs.

5.1.4 pCRF to eCRF transfer

In this scenario, clinical data are at first collected with a pCRF. This kind of documentation is in use, for example, in situations where the investigator is pressed for time or has to move between locations (e.g. emergency ward, operation theatre). In a remote data entry scenario, it is often not the investigator, but special assistance personnel who enters data from the pCRF into the eCRF. This transcription step must be quality assured. Type of personnel needed (i.e. for data entry, for data review, etc.) and criteria chosen to qualify them must be clearly defined. For using eCRF, specific training programs for investigators and assistance personal must be included.

Appropriate quality control steps have to be implemented and double data entry may be performed. pCRF transfer as well as status (arrived, reviewed, non correct, requested queries, correct, closed) must be clearly tracked. Personnel responsible for different phases of pCRF entry must be tracked as well as all the changes.
Because the investigators signature is required, he is responsible for the correct transcription of the data. Appropriate workflow support should be implemented in the EDC system.

5.2 Computerized clinical trial system

Several important requirements apply to all types of clinical data, not especially for computerized systems.

2.10 All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation and verification.

2.11 The confidentiality of records that could identify participants should be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s)

2.13 Systems with procedures that assure the quality of every aspect of the trial should be implemented.

ICH E6 Good Clinical Practice

A Quality Management (QM) system specific to data management has to be implemented, ensuring that monitoring and auditing activities with sufficient in-process quality controls are in place. Such an auditing system should be easily interrogated to identify all change points within the data record.

GCP regulates specific requirements for computerized systems.

5.5.3 When using electronic trial data handling and/or remote electronic trial data systems, the sponsor should:

a) Ensure and document that the electronic data processing system(s) conforms to the sponsor's established requirements for completeness, accuracy, reliability, and consistent intended performance (i.e. validation).
b) Maintain SOPs for using these systems.
c) Ensure that the systems are designed to permit data changes in such a way that the data changes are documented, and that there is no deletion of entered data (i.e. maintain an audit trail, data trail, edit trail).
d) Maintain a security system that prevents unauthorised access to the data.
e) Maintain a list of the individuals who are authorised to make data changes (see 4.1.5 and 4.9.3).
f) Maintain adequate backup of the data.
g) Safeguard the blinding, if any (e.g. maintain the blinding during data entry and processing).

ICH E6 Good Clinical Practice

4.9.3 Any change or correction to a CRF should be dated, initialled, and explained (if necessary) and should not obscure the original entry (i.e., an audit trail should be maintained); this applies to both written and electronic changes or corrections (see 5.18.4 (n)). Sponsors should provide guidance to investigators and/or the investigators' designated representatives on making such corrections. Sponsors should have written procedures to assure that changes or corrections in CRFs made by sponsor's designated representatives are documented, are necessary and are endorsed by the investigator. The investigator should retain records of the changes and corrections.

ICH E6 Good Clinical Practice

The SOPs listed in this document as required result from a thorough analysis of the requirements and the appropriate definition of procedures. Here national or even local adaptations must be possible. The abstraction level of the SOPs must allow the local implementation without inconsistencies.

The procedures mentioned in this document refer to processes that in the understanding of the working group have to be standardized. In this sense, they are not only recommended, but they have to be regarded as mandatory for an organization to be able to fulfil quality standards necessary for performing data management in trials supported by ECRIN.

The degree of elaboration of the specific SOPs in a trial unit performing data management in clinical trials is centre specific and should not be prescribed by ECRIN. Local interpretation of standards is possible, provided the minimal requirements are fulfilled. In clinical trial centres, often quality systems are available with their own quality systems for data management, embedded historically in other quality contexts. ECRIN should be able to audit such organizations and ascertain whether they meet the quality requirements of ECRIN. Attention should, therefore, be focused on the quality processes and their national/local implementation, not to a SOP system centrally deployed to ECRIN centres.

For this reason, this document is not aimed at establishing a SOP system with nomenclature and a numbering system but rather at giving a framework and an understanding of relevant standards and quality levels. To formalise this approach, however, it appears necessary for the SOP development to define *minimal requirements* and to distinguish them from *best practice*.

Minimal requirements for data management should be regarded as the starting level to be able to participate in common trial activities within ECRIN. In this sense, minimal requirements identify the level of minimum quality standards that could be assured to external partners wishing to make use of the network and collaborate with ECRIN. This level should be defined clearly and explicitly, so that system audits can be performed easily. Flexibility would allow each data centre to further qualify for providing more specialised or higher level services dependent on the local resources, needs and preferences. This would promote cooperation and competition within ECRIN data management expertise providers. Also some extent of specialisation would be possible (e.g. DM for trials with high dimensional data or diagnostic trials with imaging).

The procedures that must be regulated are specified in this document below. These are not titles of SOPs but rather procedures that have to be implemented within the specific local SOP system. The subitems of the list refer to minimal requirements and best practices that have to be considered in the context of each specific procedure.

The following procedures have to be regulated within ECRIN data management and appropriate SOPs have to be developed:

- **System evaluation and provider/vendor selection**

- **System installation, set up and configuration**

- **System configuration management**
  - Configuration of Audit Trail (e. g. *reason for change* optional or not?)

- **System access and profile management**

- **Change control**
  - Risk assessment of any change in the system
  - Controlled processes of making changes to the system, consisting of announcement, assessment and approval of the change.

- **System security**
  - Password policy
  - Firewall configuration
  - Physical and logical security, in particular also at the sites (EDC)
  - System controls
  - Network Security for remote access.

- **Database and communication security**
  - Encryption of data storage, data transfer
  - Electronic signature has to comply also with national regulations (EDC).

- **Data protection**
  - Handling of personally identifiable data (e. g. blinding of additionally submitted identifying data; sites should eliminate personal identifiers from source documents prior to submission)
  - Specification of minimum subject identifiers
  - Safeguarding that (future!) use of data is in accordance with informed consent
  - Regulation of access to electronic or paper based data storage
  - Particularly strict standards for genetic data
  - Secure data handling procedures
  - Use of pseudonyms / anonyms where appropriate
  - Secure cross-border data transfer.

- **Data backup and recovery**

- **Disaster system recovery**

- **Database security**

- **Data archiving**
  - Database specification
  - Data files
  - Audit trail
  - Clinical data (Open standards - vendor independent, e. g. CSV, XML, PDF, ODM from CDISC)

  - Archiving reports
  - (Scanned paper CRFs)
  - Content and variable definitions (metadata)
  - Report on data completeness at respondent and variable level
  - Secure storage and access control.

- **Business continuity**

- **Migration of data/meta-data (in case of system retirement)**

- **System validation**

- **Risk management**
  - All components of the system have to be judged according to their risk to violate GCP
  - GCP-compliance has to be guaranteed especially for high risk components
  - Maintenance of GCP-compliance even after updates or other changes to the system.

- **Periodic review/audits**

- **Safeguard of blinding**

- **Help desk**

In particular, data protection and data security may be liable to national regulations and should be reflected in the SOP system accordingly.

In general, it is proposed that standards relevant to all these procedures would be recommended and adopted by all ECRIN members. To facilitate this adoption, specific SOPs and process control documents will be defined.

5.3 Training and qualification

5.6.1 The sponsor is responsible for selecting the investigator(s)/institution(s). Each investigator should be qualified by training and experience and should have adequate resources.

ICH E6 Good Clinical Practice

2.8 Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective task(s).

ICH E6 Good Clinical Practice

5.5.1 The sponsor should utilise appropriately qualified individuals to supervise the overall conduct of the trial, to handle the data, to verify the data, to conduct the statistical analyses, and to prepare the trial reports.

ICH E6 Good Clinical Practice

It is proposed that standardised job descriptions, including mandatory and desirable qualifications and experience, will be created and distributed to all ECRIN members.

Different roles for personnel must be defined. In data management the following roles are required:

- Clinical investigator
- Project coordinator
- Database manager
- Network administrator
- System administrator
- Statistician
- Data manager
- Data entry personnel
- Data monitor
- Quality manager.

Definitions to include:

- Roles needed in a specific project
- Qualification and skills needed for each person involved
- Content of general training of the staff involved in data management
- Content of training for specific projects
- Documentation of general and specific training
- Internal SOPs giving rules on training (i.e. how often, where, reason for, request)
- Policy on staff recruitment (i.e. use of external expert, only employee).

As part of wider ECRIN training programmes, data management modules will be included. Recognising the multi-site nature of ECRIN, such training may need to be developed in a distributed web-based e-learning way.

For all roles involved in data management, the minimum training requirements for new staff include:

- Introduction
- GCP (Good Clinical Practice) training
- Risk assessment
- Data Protection Act and Institutes own Data Protection Policy
- Institutes own Information Security Policy
- Good Data Management Practices.

Desirable qualifications:

For data managers a degree in life sciences (e.g. medicine, biology), natural science (e.g. computing, statistics, mathematics) or another adequate qualification and education/training in data management is desirable, whereas for data-entry personnel thorough instruction by the data manager is sufficient.

Courses organised by the Association for Clinical Data Management include:

- Introductory, intermediate and advanced clinical data management
- Computer systems validation in clinical research
- Postgraduate qualification in clinical data management.

Courses that include vocational and competence-based qualifications are currently being jointly developed by the Association for Clinical Data Management and the Association for the British Pharmaceutical Industry.

5.4 Data management procedures

2.6 A trial should be conducted in compliance with the protocol that has received prior institutional review board (IRB)/independent ethics committee (IEC) approval/favourable opinion.

ICH E6 Good Clinical Practice

5.4.1 (e)CRF design

Acquisition or collection of data can be achieved through paper or electronically. GCP uses the term “case report form” to refer to this process.

1.11 Case Report Form (CRF)
A printed, optical, or electronic document designed to record all of the protocol-required information to be reported to the sponsor on each trial subject.

ICH E6 Good Clinical Practice

According to GCP a CRF covers both paper-based data collection (pCRF) and electronic data collection (eCRF). Because there are different processes describing the use of paper-based and electronic tools for data capture, different use cases were created to analyse differences in requirements.

For this purpose, different scenarios are presented in appendix III covering a range from paper-based data collection to full electronic data capture. In practice, often a mixture of paper-based and electronic data collection is used. The following requirements/best practices are both applicable to pCRF and eCRF:

- ✓ Design the (e)CRF to collect only the data specified by the protocol
- ✓ Document the process of (e)CRF design, review and versioning
- ✓ Proofread the paper-CRF

- ✓ Proofread documentability of flow of patients in relation to CONSORT flow and checklist
- ✓ Deliver (e)CRF to sites prior to enrolment
- ✓ Ensure the (e)CRF does not duplicate data or calculated results unnecessarily
- ✓ Avoid redundant questions, if not for validation purposes (then using different measurement means)
- ✓ (amended) CRF pages reflect the current and correct version identifier
- ✓ Ensure that primary safety and efficacy variables drive the CRF development
- ✓ Proofread relevant topics in relation to key protocol variables (e.g. inclusion/exclusion criteria, endpoints, serious adverse events)
- ✓ Be familiar with validated questions/standard instruments and maintain the integrity of validated questionnaires (e.g. quality of life questionnaires)
- ✓ Apply common documentation principles for data items (prefer coding, numbering of items, minimize referential questions, specify type of missing data, don't mix raw and calculated data, prefer positive formulated questions, define complete answer categories, etc.)
- ✓ Training of clinical site personnel and investigators in the use of (e)CRF

- ☺ Design the (e)CRF considering the workflow of trial procedures and organizational aspects (e.g. who will enter data where)
- ☺ Design the (e)CRF together with those responsible for data analysis, data managers and end-users
- ☺ Implement a library of forms
- ☺ Implement a library of data items (including metadata, e.g. ranges, units, site contact details)
- ☺ CRFs should be reviewed against the protocol, end-user expectations and CRF design best practices and comments provided in a timely constructive manner
- ☺ CRF should be in the correct format and translated into the appropriate language(s) – verification of translation may be required
- ☺ Ensure that the CRF is divided into appropriate sections with simple and clear instructions for completion
- ☺ Amended CRF conforms to requested amendments and/or revised protocol
- ☺ CRF page numbering and version information is updated to reflect current status of document
- ☺ A site will only have access to data collection systems once all relevant paperwork has been completed; including ethical and research approvals, contracts, site initiation and (e)CRF training.

The following SOPs are needed for data acquisition:

- (e)CRF-design
- (annotated) CRF-review (if prepared externally)
- (e)CRF approval/release
- (e)CRF amendment
- (e)CRF version control

- (e)CRF library management
- (e)CRF training.

5.4.2 Database validation, programming and standards

The validation process covers the validation of the clinical data management system (CDMS) in general (system validation) and the validation of the trial-specific eCRF and database. The use of a validated system is a prerequisite for GCP-compliant data management.

5.5.3 When using electronic trial data handling and/or remote electronic trial data systems, the sponsor should
a) Ensure and document that the electronic data processing system(s) conforms to the sponsors established requirements for completeness, accuracy, reliability, and consistent intended performance (i.e. validation).

ICH E6 Good Clinical Practice

Management of the Database Management System (DBMS)

Management issues related to DBMS are:

- Ensure the security of the DBMS and clinical trial data
  - o Back-up and disaster recovery
- Maintain the DBMS according to user requirements
  - o Appropriate levels of access
  - o Upgrade and changes
- Provide training in the use of DMBS
- Ensure system templates are maintained
- Time and vendor independent solution and format of archiving and export of data.

Trial-specific design and validation

Database validation should be based on aggregated validation of all design elements of a database. Database design elements are CRFs, profile documents, ACL (Access Control List), programming code and validation routine, configuration documents (replication and system settings), user group definitions, template information and encryption keys.

For trial-specific design and validation the following requirements/best practices should be taken into consideration:

- ✓ Define test methodology with a trial-specific test plan
- ✓ Review of programmed code/validation of programmed modules
- ✓ Test data should be used to appropriately test the database set-up and data entry screens
- ✓ Test data are entered accurately and kept separate from live data
- ✓ Data validation checks are programmed as specification

- ✓ Data validation checks are run on test data and the checks raised are compared with those expected for the data entered
- ✓ Data validation programs (scripts) should be checked, tested and documented
- ✓ Any issues regarding database set-up or data entry will be identified by checking against expected results, dealt with and documented

- ☺ Follow coding/programming principles for database design/ programming
- ☺ Creation of library of reusable validated code
- ☺ Documentation principles for standards/certification to be met by the software used
- ☺ The process of database design and checks programming should be reviewed at any step for assuring adherence to protocol
- ☺ Develop user guides including documentation.

The following SOPs are needed for database validation, programming and standards:

- Database design and programming
- Validation of programmed modules/applications
- Database review/validation (trial specific).

5.4.3 Randomization and blinding

4.7 Randomization procedures and unblinding. The investigator should follow the trial's randomization procedures, if any, and should ensure that the code is broken only in accordance with the protocol. If the trial is blinded, the investigator should promptly document and explain to the sponsor any premature unblinding (e.g., accidental unblinding, adverse event) of the investigational product(s).

ICH E6 Good Clinical Practice

5.5.3 When using electronic trial data .... g) safeguard the blinding, if any (e.g. maintain the blinding during data entry and processing).

ICH E6 Good Clinical Practice

5.5.5 The sponsor should use an unambiguous subject identification code (see 1.58) that allows identification of all the data reported for each subject.

ICH E6 Good Clinical Practice

Data management may be involved in the randomization process via providing an electronic randomization service to the sites or by performing telephone or fax randomization using in-house lists or programs.

- ✓ The randomization system must be validated including all procedures interfacing with biometrics. System controls should monitor the randomization process regularly. Standards are necessary to handle errors or problems in the randomization process as this is crucial to the overall quality of the trial.
- ✓ The specification of the randomization design is the responsibility of the study statistician. They will communicate this specification in writing a Randomization Specification, including whether mechanisms have been put into place securing balanced numbers of patients in each randomization group.
- ✓ The application programmer will be responsible for coding the specification, and producing test runs to validate the randomization code.
- ✓ Unblinding of a treatment allocation (if applicable) due to safety concerns must be requested by the appropriate person and the data, time and reason for unblinding must be recorded.
- ✓ Throughout the conduct of the trial no persons will have access to unblinded data (exception may be possible for Data Monitoring Committee). Statistician should be blinded until final analyses are agreed.
- ✓ At trial analysis, the data will be unblinded to the rest of the research team only when final analyses have been formally conducted in accordance with the agreed Statistical Analysis Plan.

SOPs needed for randomization/blinding:

- Randomization
- Unblinding.

5.4.4 Data entry, data processing and data validation

5.1.3 Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly.

ICH E6 Good Clinical Practice

5.5.3 When using electronic trial data
d) Maintain a security system that prevents unauthorized access to the data
e) Maintain a list of individuals who are authorized to make data changes (see 4.1.5 and 4.9.3)
f) Maintain adequate backup of the data

5.5.4 If data are transformed during processing, it should always be possible to compare the original data and observations with the processed data.

ICH E6 Good Clinical Practice

Data entry

The data entry process should be defined for the specific trial and specified in a **Data Management Plan**. For transcription from pCRF to eCRF different procedures are used:

- double data entry (one person)
- double data entry (two persons)
- single entry with second look
- single data entry with reading aloud
- single data entry with source verification
-

Double data entry is not required by regulations but "good practice". The data entry process should be chosen based on the skills of the personnel, the resources in the project and the reflected evaluation of key variables.

4.9.1 The investigator should ensure the accuracy, completeness, legibility, and timeliness of the data reported to the sponsor in the CRFs and in all required reports.

ICH E6 Good Clinical Practice

4.9.2 Data reported on the CRF, that are derived from source documents, should be consistent with the source documents or the discrepancies should be explained.

ICH E6 Good Clinical Practice

Raw data:
Data as originally collected. Distinct from derived. Raw data includes records of original observations, measurements, and activities.

CDISC Clinical Research Glossary, Version 6

Raw data should be collected wherever appropriate. The site should not have to derive or calculate any values. Special care should be given to the problem of differing laboratory ranges and/or units.

- ✓ Format of data to be received from external systems agreed and standardised
- ✓ Data entry according to agreed instructions
- ✓ List of authorized persons for data entry
- ✓ User training with data entry instructions/guidelines necessary
- ✓ Documentation of data receipt
- ✓ Audit trail for data entry
- ✓ Data received should be checked and any transfer problems identified
- ✓ Ensure blinding of information submitted to the data centre with regard to subject identifying information

- ☺ Tracking of pCRF-pages before data entry
- ☺ Time-lines for data entry.

SOPs needed for data receiving and data entry:

- Data receipt and storage of pCRFs during data entry
- Data entry
- User training

Data processing

✓ All transactions to the database (insert, update, delete) must have a clear and complete audit trail
✓ Data only accessible to authorised personnel
✓ Site staff only access to data of their site
✓ Data handler familiar with GCP and will keep data secure and confidential at all times
✓ Coding performed using appropriate dictionaries
✓ Where autocoding is not possible, manual coding is performed

☺ Audit trail searchable and capable of producing audit trail reports
☺ Coding conventions should be observed to ensure consistent coding within and between studies
☺ Use of an autoencoder and synonym list where possible.

SOPs needed for data processing:

- Medical coding.

Data validation

The level of quality controls applied to data must be transparent. Any procedure involved in data cleaning, performed manually or automatically by validation check programming, has to be predefined in a **Data Management** and **Data Validation Plan,** preferably outlined in the protocol. In case of a change of validation rules during the conduct of a trial revalidation of all data might be necessary, requiring an additional and often labour-intensive step.

✓ Data quality checks carried out according to agreed instructions and GCP and regulatory requirements
✓ Manual checks (i.e. visual checks of CRFs with manual review of the data, e.g. medical consistency checks, lab data pointing to an AE)
✓ Computerised checks (e.g. immediate checks during data entry or checks to be run in batches, e.g. at the end of a visit module or at the end of the CRF)
✓ Checking of missing, illogical and inconsistent data
✓ Complete documentation of data checks
✓ Errors reported to the appropriate person for resolution
✓ Final data checking.

SOPs needed for data validation:

- Data validation plan (quality assurance actions applied or data validation)
- Data management plan (data management activities and data standards used).

Query management

Queries are defined by the CDISC glossary:

> Query:
> A request for clarification on a data item collected for a clinical trial; specifically a request from a sponsor or sponsor's representative to an investigator to resolve an error or inconsistency discovered during data review.
>
> Query management:
> Ongoing process of data review, discrepancy generation, and resolving errors and inconsistencies that arise in the entry and transcription of clinical trial data.
>
> Query resolution:
> The closure of a query usually based on information contained in a data clarification.
>
> ***CDISC Clinical Research Glossary, Version 6***

Before locking the database, there should be an agreed list of validation checks, which can be performed on the data for checking of consistency, etc. Queries as part of data analysis are not considered in this document.

- ✓ Queries should be created in accordance with customer requirements and documented procedures (data review guidelines and data validation plan)
- ✓ Defined procedure for self evident corrections performed by data management staff
- ✓ Query resolution tracked and action taken within agreed time-scales
- ✓ Action taken on queries is appropriate and edits are documented
- ✓ All transactions to the database (insert, update, delete) must have a clear and complete audit trail
- ✓ Adequate SOPs and working instructions for data changes
- ✓ Take into consideration trial amendments, which may have consequences on the CRF

- ☺ Ensure no duplication of queries
- ☺ Single checks with all variables, complicated checks with critical variables
- ☺ Queries are issued to sites within agreed time-scales
- ☺ Queries should have response within agreed time-scales
- ☺ Reports on query management.

SOPs needed for query management:

- Query management
- Database and user audit (may be dealt with by other QA measures).

Data Management Plan

An important element in data management is the **Data Management Plan (DMP).** A Data Management Plan should describe and define all data management activities for a study. The DMP should define the procedures that describe how the data will be managed and to what standards. The DMP could be used to document study-specific deviations from standard SOPs. Procedures specified already in SOPs does not have to be in the DMP.

Essential components of a DMP are:

- Map of file server arrangements, etc (alternatively with a reference to other documents)
- Details of study personnel involved in the study and data access roles assigned to each
- A complete set of finalised case report forms (CRFs) and amendments
- Database design (alternatively with a reference to other documents).
  - Software, hardware and database location
  - Detailed description of database structure (Data Dictionary)
  - Detailed description of data entry system
- Data entry procedure
  - Methods of data collection – paper CRF, electronic devices, etc.
  - Type of data entry – double or single data entry with checking
  - Data preparation before entry onto electronic system
- Data Query Rules
  - Automated checks should be specified in sufficient detail to enable set up of data entry screens and validation programs. Checks that can be done automatically during or after data entry should be clearly identified.
  - Data flow and tracking to ensure optimal data completion and to facilitate reporting
- Query Handling
  - How queries will be tracked
  - Expected resolution time for data queries
  - Who is responsible for making required changes to the data
  - Who is responsible for ensuring all queries are resolved before data is locked for analysis
  - List of agreed queries
- Quality Assurance Plan should include:
  - Audit trail checks
  - Sample checks of critical data
- Data review checks to support monitoring
- Training plan and log for data entry systems if required
- Electronic data transfer rules
- Back-up and recovery procedures (alternatively with a reference to other documents).
- Archiving and security arrangements
- Reporting Progress

5.4.5 Database lock/database archiving

Database lock/unlock

This controlled procedure to freeze the data in a specific status usually is performed preventing write/edit access to most users of the system, thus securing the data from any further changes. Unlocking of the database should be a very strictly controlled and documented step.

- ✓ Procedure for lock/unlock must be clearly defined
- ✓ All data must have been received prior to database lock
- ✓ All cleaning procedures must have been completed
- ✓ All queries should have been resolved
- ✓ External data (e.g. safety database, lab data) must have been *reconciled*
- ✓ Final consistency check of database (also with statistical methods)
- ✓ Conditions for unlock should be defined

- ☺ Coding must have been reviewed
- ☺ A database audit might be useful (documenting error rate).

SOPs needed for database lock/unlock:

- Database lock/database unlock.

Data archiving

- ✓ Definition of CDMS export format
- ✓ Patient identifiers should not be archived with clinical and outcome data
- ✓ Data should be archived securely (e.g. locked rooms and fire-proof cupboards, safe area, protected and controlled access for authorized staff only)
- ✓ Definition of procedures for data archiving and data access (database access, user access, system controls)
- ✓ Documentation of access to study archive
- ✓ Data should be archived for as long as specified by regulations, funding body and/or sponsor
- ✓ Data archiving of unprocessed data after database lock
- ✓ Complete documentation in study archive (e.g. database structure and programming, final data sets, audit trail, originals of study documents, documentation of non-compliance to SOPs and Working Instructions, documentation of database lock/unlock, centre specific data copies)

- ☺ Use of standardised formats for archiving (e.g. ASCII, PDF, XML, CDISC ODM, FDA approved SAS format)
- ☺ Adequate use of safety copies (e.g. scanning of paper documents)
- ☺ Electronic data should be decrypted before archiving. In case of encryption of data time-independent decryption methods should be in

place. As an alternative to encryption archiving in research data archives could be used.

SOPs needed for data archiving:

- Database archiving
- Archiving of essential documents
- Archiving (e)CRF.

5.4.6 Data quality

5.5.4 If data are transformed during processing, it should always be possible to compare the original data and observations with the processed data.

ICH *E6 Good Clinical Practice*

Procedures are required for measuring and assuring data quality. Quantitative methods should be used. Statistical methods or industry standards of determining sample sizes for quality controls should be used to define the sample size of data to be reviewed. Error rates should be measured (CRF-to-database inspections); alternatively strict in-process quality processes may be used to assure data quality.

For data quality management, the following requirements/best practices should be taken into consideration:

- ✓ Data quality checks should be carried out according to agreed instructions and with quantitative methods (at least one quality inspection). If validated working processes for data transformation have been implemented, the procedure may be simplified
- ✓ Quality checks should be fully documented
- ✓ Errors should be reported to the appropriate persons for resolution (e.g. corrective actions in data handling report)

- ☺ Comparison of final data with source data, if there are many data processing steps
- ☺ 100% check of primary target variables and other essential data
- ☺ Performance of centre-specific plausibility checks (e.g. monitor aggregate data by site)
- ☺ Performance and evaluation of data quality impact on analysis
- ☺ Use of quality adapted criteria for performing monitoring/triggering site audits
- ☺ Assessment of compliance of DM process with GCP, rules and regulations
- ☺ Statistical methods should be used to assess and evaluate data quality (e.g. appropriate statistical sampling to measure data quality); measures to analyse possible problems and irregularities should cover e.g. multivariate analysis of possible outlier candidates, conspicuous data patterns, preferred numerical sequences, accumulation of values close to defined limits.

SOPs needed for data quality:

- Measuring data quality
- Data quality acceptability criteria
- Monitoring data quality.

5.4.7 Reproducible data analysis

All analyses must be carried out in a form where direct reference (reading) is based on verified and anonymized copies of the “locked database”. A collection of scripts should be presented as part of the verification of all published results, whereas scripts and queries from daily analysis work should be kept in a structured way available for external review. There is no demand for a complete audit trail of all queries to the “locked database”.

SOPs required:

- Study specific principles of data analysis and documentation of analysis.

5.5 Safety data management

A safety database should be usually provided for the collection of Serious Adverse Events (SAEs).

SAE reconciliation

If a dedicated safety database has been set up, the data in the clinical database and those in the safety database have to be reconciled. For regulatory purposes and to meet obligations to report SAEs, it should be guaranteed that data in both databases are identical.

SOPs are needed to define e.g. the point in time for SAE reconciliation (SAEs might be reported also after database closure):

- SAE reconciliation.

5.6 Dictionary management

Coding of Adverse Events and medication should be implemented. For the coding of Adverse Events (AEs) MedDRA is the accepted and expected standard. Auto-encoding systems should be used when available. It should be established, how to handle changed versions of dictionaries during the term of a project.

Manual changes of public dictionaries are not allowed. If necessary, a private dictionary as an interface to a public dictionary should be implemented (synonym list).

SOPs needed for dictionary management:

- Coding with MedDRA (Coding Guidelines)
- Medication coding
- Dictionary maintenance
- Dictionary version control
- Auto-encoder configuration.

5.7 Interfaces

Laboratory data

Management of laboratory data and normal ranges (e.g., detection of possible Adverse Events in clinically significant outliers) should be considered already in the design of the clinical trial database (see 5.4.1). For import of lab data from a lab system into the CDMS, it is recommended to take the CDISC Lab standard into consideration.

The issue of different lab ranges and units between sites and countries is a well-known problem. Software should be used for automated import of data of different laboratories. This presumes, that the software for laboratory data import supports detailed configuration for multisite lab data, including generic and highly configurable interfaces for different data formats and options for quality assurance of the imported data. Another option could be to impose a maximum number (e.g. 2 units) of different units for a lab parameter and let the sites do the conversions themselves.

Data exchange standards

The use of data interchange standards is especially relevant in multicenter global trials, where the pooling of data delivered by different EDC systems are an issue. This domain is being elaborated by the CDISC consortium. For data interchange, the ODM standard was developed. Therefore, it is recommended to implement procedures to export study data from CDMS in ODM. SAS-format could be another option.

Many systems for clinical data management have already implemented the ODM data model. The system evaluation and implementation planning activities within ECRIN should give great importance to this issue, which will not be discussed here.

5.8 Important documents for data management

Many documents are produced within a clinical trial. A common set of specific documents would greatly improve harmonisation and interoperability. Several important documents to support compliant data management were identified:

- Study database validation plan, test plans
- Validation report
- Data management plan
- Annotated CRF
- Blank (unmarked) copy of CRF
- Mock Up CRF (optionally, for usability test)

- Edit specifications
- Data entry guidelines
- Site qualification, signature sheets
- Access control list
- (e)CRF training documentation
- Data validation plan
- Data review plan (for medical checks e.g. of medical consistency and AEs)
- Data handling report
- Database audit report
- Database lock documentation
- List of variables and reference values.

5.5.6 The sponsor, or other owners of the data, should retain all of the sponsor specific essential documents pertaining to the trial (see 8. Essential Documents for the Conduct of a Clinical Trial).

ICH E6 Good Clinical Practice

Essential documents have been defined by ICH GCP. All documents with regulatory character (i. e. any document that might be relevant for an audit or an inspection) should be subjected to a version control system. It makes sense to have SOPs for the whole life cycle of such documents, be it electronic or paper. Special care should be provided in respect of a procedure regulating version control.

SOP necessary:

- Data management documentation (incorporating standard procedures on the creation of data management relevant documents, e.g. Data Management Plan).

5.9 Electronic data capture

General statements

The use of electronic data capture and other primarily electronic processes in clinical trials can have some advantages over paper-based clinical trials (see Appendix III: Scenarios).

Some major advantages are:

- automated data edit checks during data entry
- rapid access to previously entered data, administrative, clinical, biological, etc.
- immediate availability of data for quality assurance and for reporting to the sponsor (e. g. enrolment status in big trials, or number of forms with uncorrected errors, etc. should be readily available)
- easier collection of patient reported outcome data, with higher consistency
- integration of information from areas outside clinical research (e. g. basic research, translational research) is facilitated provided the eCRF system provides adequate interfaces to other systems

- monitoring is supported when monitors are given online access to the database
- reduction of data needing clarification as interactive data entry at the investigative site allows immediate data checks and rapid data collection
- earlier occurrence of database lock.

It includes also some drawbacks:

- time and resources needed to train users with respect to eCRF
- data entry is shifted from the centralised data handling unit to the investigative site. This may imply that data entry may be performed by non-specialised personnel, making a secondary check necessary by another person
- the advantages of rapidity are highly dependent on the time available for data entry. If members of the investigator team have not the required availability, there may be a delay between data collection and data entry
- two situations have to be distinguished: a) direct data entry into eCRF or b) data collection primarily done on paper and then entered in the eCRF (see appendix III: scenarios). In the second situation the workload is significantly increased and data management steps are prone to errors similar to the classical paper approach. In general, the use of eCRF mixes the two situations, represented in proportions that vary from trial to trial.

Some critical issues in the set up of EDC trials:

- the system used must satisfy regulatory requirements such as the audit trail
- the database must represent the source data in an auditable way (only for eSource; when paper CRF are filled in the eCRF as a second step, then the paper CRF can be part of the source data – see Appendix III: Scenarios)
- the privacy of the participant's data must be retained at the site, especially if eSource data are used
- the trustworthiness of the host providing the system must be clear, especially if web based applications are used, accessing one central database (the site must be in control of the data - this could require a trusted third party hosting the database and **not the sponsor**)
- appropriate interfaces have to be developed.

The implementation of electronic data capture (EDC) trials requires additional tasks and procedures, including the appropriate quality assurance. The following procedures will have to be considered in SOPs:

System installation, set up and configuration

- ✓ Ensure time synchronization within the system. Sites should not be able to change the system's time stamp (audit trail!)
- ✓ Complete system validation prior to trial implementation (no *retrospective* validation). Validation should be performed according to 5.2 taking into consideration specific aspects related to eCRF use.

- ✓ Qualify sites for participation
- ✓ Ensure system security and access control
- ✓ Provide Help Desk and Hot Line Help
- ✓ Perform system training

- ☺ Ensure compliance with FDA requirements and guidances (21 CFR 11, CSUCT) if FDA-compliance is needed
- ☺ Ensure system availability and usability
- ☺ Try to avoid last minute changes or modifications, due to bad project management
- ☺ Ensure that audit trail is easily retrievable and readable to authorized users
- ☺ Define processes for data integration and transfer
- ☺ Perform change control on all system changes, also on "user-configurable" settings
- ☺ Implement an audit trail for metadata changes
- ☺ Automate the generation of reports.

eCRF design

- ☺ Use libraries with procedures concerning library management ("library custodian")
- ☺ Separate eCRF design from edit checks programming (if supported by system, if personnel available)
- ☺ Provide a simplified pCRF for discussion.

eCRF validation

- ✓ Document any errors for tracking purposes
- ✓ The test environment and the production environment have to be identical (configuration management!), [so in general three instances of the system are recommended: *Test, Development, Production*]

- ☺ A V-model oriented procedure is recommended
- ☺ Define the testing methodology within a trial specific *Test Plan* covering scope of test, item pass/fail criteria, etc. Software risk issues should be evaluated and approaches/strategies developed to cope with them.

eCRF acceptance testing

- ✓ Test the eCRF's usability involving key site personnel and monitoring
- ✓ Document acceptance (i.e. let key persons sign their acceptance of the eCRF)

eCRF training

- ✓ Document (e)CRF training
- ✓ Document training on protocol
- ✓ Provide (e)CRF completion instructions

eCRF release

- ✓ Make clear to any user whether they are working on a test eCRF or whether the "real trial" has been opened (e.g. by locking the test database and resetting all accounts)
- ✓ Inform all users about the trial deployment
- ✓ Restrict access to trained users

eCRF amendment

- ✓ Submission of change requests only by designated people
- ✓ Accumulation of change requests to minimize amendments
- ✓ A risk analysis is necessary before any amendment
- ✓ Notify all users about changes
- ✓ In the case of significant changes, the need for a re-training should be evaluated

☺ First, test any amendment in the test environment, following a test specification that has been developed based on the risk analysis

Site qualification

The sites should be able to participate in the technical infrastructure of the eTrial. Thus, the data management team has to assure and document that the sites meet certain requirements. Technical issues such as VPN, firewalls, connectivity, previous experience with EDC trials etc. have to be required elements. Only qualified sites should be able to get login accounts for the EDC trial, after having performed the necessary eCRF training.

The following procedures should be followed:

- Provide sites with requirements (hardware, operational system, applications, antivirus software, system security)
- Ensure site adopts specific relevant SOPs
- Ensure site personnel have been trained in system use
- Make sure site meets technical and administrative requirements
- Collect a signature (prior to granting access to any user) that reflects the FDA's requirements if necessary
- Document the contact information of technical or administrative individuals at the sites
- If installation is needed at the site, perform validation of installation (e.g. installation qualification; preferably in automated manner).

5.10 Quality management system

Quality is a measure of the ability of a product, process, or service to satisfy stated or implied needs. In data management, quality may apply to data and processes.
The core element of a data management quality system is the development and application of standard operating procedures. The SOPs should amongst other things

- define responsibilities

- specify records to be established and maintained
- specify methods and procedures

The SOP system should be used coupled with investigator's supervision of the trial's conduct and with sponsor's monitoring (quality control). Independent auditing activities performed by the sponsor or by regulatory bodies (inspections) is the major activity of quality assurance. Knowledge gained within the performance of quality assurance which is systematically evaluated and analysed can be used for quality improvement.
The quality system in data management is embedded in an overall quality system. Specific issues mainly occur in reference to quality control and internal, independent audits.

6. Multinational aspects

In multinational trials, several very specific issues that are associated with regulatory, design, language, and cultural aspects need to be addressed. These multinational aspects have been considered within the European Union where some degree of regulation harmonisation and common procedures of data management for trials across countries exist. A multinational trial involving non-European countries (e.g. Arabic, African, or Chinese) could increase the difficulties associated with its performance due to the substantial level of country heterogeneity. In this case, a lot of consideration and preparation is required. In its starting phase, however, ECRIN will focus on multinational European trials.

Different legislation may apply in different countries, amongst other things concerning privacy, data encryption, transfer of data across borders, etc. However the publication of robust standards and processes may serve to address such issues.

A multinational trial can be considered as a pool of multi-centre trials performed according to the same protocol. The differences in population, culture, nomenclature, and medical practice can be possible causes of greater bias and variability than in single country trials. Of course, also in single country multi-centre trials variability and bias can differ, dependent on the homogeneity of population, culture etc. Some of these possible sources of bias can be controlled by the trial design, using stratification or other statistical methods, some of them addressed in the analysis phase. Others may require investigator training, e.g. to overcome differences in diagnosis culture or in the assessment of efficacy and safety parameters, due to cultural aspects (especially in the reporting of Adverse Events). The definition of standardised trial processes is important to ensure standardisation of trials in this context.

Modern concepts for the validated transfer of outcome measure instruments, in particular quality of life patient diaries, into other languages/cultures make use of standardised translations and back translations, to assure both conceptual and semantic equivalence. Translation will always increase the trial budget, unless officially verified versions of scales and indexes are used[1].

[1] Examples of such translations are EuroQol Eq-5d (see www.euroqol.org)

Before start of a multinational clinical trial, it has to be decided in which language

- trial documents
- CRFs
- SAE-forms

have to be provided and used. Persons, responsible for translation, should be determined, and a time-table should be agreed upon. The necessary details should be specified in a document (e.g. contract).

Data management is involved in respect of Case Report Form translation issues, technical issues (e.g. multi language electronic CRFs will be more hard to cope with if pure web-based technology is used) and system configuration (e.g., differing regulation on data encryption, which has to be implemented accordingly if EDC is used). It has to be specified, in which language data should be entered into data fields. This is particularly relevant for free text data entry into SAEs and description of SAEs. If it is permissible to enter data in a language different from English, it has to be decided who has to perform translation and when, and whether any references are needed (back-translation). In any case, it must be specified, whether an independent retranslation is necessary to achieve adequate data quality. In particular with respect to AE reporting, free text in national languages may cause a problem for data handling, e.g., if auto-coding engines have to be configured adequately (This problem can be solved by using multilingual dictionary -- MedDRA, WhoDrug. The application of such a solution would require a coding team per language/country, clear and standardized coding rules, and a specific training to apply these rules. A specific SOP to define these rules could be proposed in the framework of ECRIN.) The original text will usually occur in national language; therefore, a validated translation will be necessary.

For multilingual application (not only in English), the validation process should involve at least one person of each language/country. A minimal validation plan is necessary to ensure the CRF and message translations are correctly coded for each language. The issues raised by translation of the text of the eCRF are safer if the text is managed as a database, independent of the eCRF code.

It must be determined if the applications used can handle the different characters used for certain languages (e. g. Greek in Greece, Cyrillic in Belarus, Ukraine, Bulgaria, Serbia).
The methods used for data collection can have advantages and disadvantages in a multinational setting. e.g., a central CRF design may not be desirable or applicable in all countries. Patient diaries always need translation. If instead phone interviews could be used, this would possibly have advantages in some settings. Informed consent and patient reported outcomes (quality of life instruments) in any case will have to be translated and cultural differences also considered.

Lastly, the use of local laboratories will be the usual procedure in ECRIN trials. The pooling of data from different labs requires quality assurance steps during data collection and import, making use of normalised data, units and ranges and assuring the reproducibility and repeatability of those data.

7. Perspectives for ECRIN data centres

To enable the efficient conduct of multinational clinical trials in ECRIN, the concept of ECRIN data centres will be developed. These data centres will offer data management services and the corresponding software tools to the ECRIN community. Thus, cooperative international data management will be supported by these ECRIN data centres. In the ongoing ECRIN-project ("European Clinical Research Infrastructures Network and biotherapy facilities: preparation phase for the infrastructure (ECRIN-PPI)"), a framework and the requirements and conditions for becoming an ECRIN data centre will be specified. It is the objective of ECRIN-PPI to implement at least one prototype ECRIN data centre before the end of this project and to make appropriate decisions for the optimal number and funding of further ECRIN data centres thereafter.

ECRIN data centres will not be constructed from scratch but be based upon existing resources and competencies available within the ECRIN network. It is planned, that clinical trial units/centres in ECRIN will have to undergo an evaluation and certification process to demonstrate their ability for providing GCP-compliant data management for multicenter trials. Successful centres will be qualified to become ECRIN data centres. The requirements for this evaluation process will be developed based on the information contained in this document. Details of the process of implementation of ECRIN data centres will be worked out by an IT-coordination group implemented at the Coordination Centre for Clinical Trials of the Heinrich-Heine-University, Düsseldorf, Germany and with strong involvement of the Working Group members of ECRIN. This group will define user requirements, software specifications, quality targets, audit procedures, software evaluation and validation processes.

There exist already a number of centres in ECRIN, experienced in GCP-compliant data management. These centres or new emerging data centres may apply to become an ECRIN data centre in accordance with the framework and conditions specified by the IT-coordination group. Therefore, application documents will be provided. Calls and evaluations are performed by the ECRIN IT-coordination group via audit procedures and can lead to approval and certification with the status of an ECRIN data centre. The process to qualify for being an ECRIN data centre is an important part of ECRIN's approach to conduct trials. Therefore, ECRIN may cover some of the costs associated with the qualification procedure (e.g. audits); however, no costs associated with running a data centre (e.g. hardware, software, DM personnel). It should be stressed, that prerequisite for becoming an ECRIN data centre, apart from compliance with this document and the requirements and conditions to be specified, will be the commitment to support multinational clinical trials according to service conditions specified by ECRIN.

8. Appendices

Appendix I: Terms and abbreviations

AE – Adverse Event
CDISC – Clinical Data Interchange Standards Consortium
CDMS – Clinical Data Management System
CONSORT – Consolidated Standard of Reporting Trials
CRF – Case Report Form
CSUCT – Computerized Systems Used In Clinical Trials
CSV – Comma Separated Values
DBMS – Database Management System
DDE – double data entry
DM – Data Management
DMP – Data Management Plan
EC – European Commission
ECRIN – European Clinical Research Infrastructures Network
eCRF– electronic Case Report Form
EDC – Electronic Data Capture
EDP Electronic Data Processing
EMEA – European Medicines Agency
FDA – Food and Drug Administration, US Department of Health and Human Services
GCP – Good Clinical Practice
ICH – International Conference on Harmonisation
IVR – Interactive Voice Response
MedDRA – Medical Dictionary for Regulatory Activities
NIHR -- National Institute for Health Research
ODM – Operational Data Model (CDISC)
pCRF– paper Case Report Form
PDF – Portable Document Format
QM – Quality Mangement
RDE – Remote Data Entry
SAE – Serious Adverse Event
SAE reconciliation – the process of investigating clinical data and safety data in order to detect discrepancies (e.g. review of CRF and SAE data) and the process of resolving those discrepancies
SAS – Statistical Analysis System
SDTM – Study Data Tabulation Model (CDISC)
SOP – Standard Operating Procedure
STATA – Data Analysis and Statistical Software
SUSAR – Suspected Unexpected Serious Adverse Reaction
VPN – Virtual Private Network
XML – eXtensible Markup Language

Useful Glossaries

- Glossary in “Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95)”
- CDISC Clinical Research Glossary, Version 6.0
- CDISC Acronyms, Abbreviations, and Initials, Version 6.0, Applied Clinical Trials, Dec. 2007, 12-40.

Appendix II: References

Regulatory documents

- ICH Topic E6: Guideline for Good Clinical Practice Guideline, Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95), EMEA, January 1997
- Directive 2001/20/EC of the European Parliament and of the Council on the approximation of the laws, regulations and administrative provisions of the Member states relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. Official Journal of the European Commission of 4 April 2001, No. L 121 p. 34
- Directive 2005/28/EC of the European Commission laying down principles and detailed guidelines for good clinical practice as regards investigational medicinal products for human use, as well as the requirements for authorisation of the manufacturing or importation of such products. Official Journal of the European Commission of 9 May 2005, No. L 91/13-L91/19
- Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Official Journal of the European Communities of 23 November 1995, No L. 281 p. 31.

US-based documents

- FDA, Guidance for Industry. 21 CFR Part 11, Electronic Records; Electronic Signatures – Scope and Application (August 2003)
- FDA, Guidance for Industry. Computerized Systems Used in Clinical Investigations (CSUCT) (May 2007)

Draft Documents

- EMEA. Reflection on expectations for electronic source documents used in clinical trials. London, 17 October 2007 (draft).

Documents with normative character

- World Medical Association: Declaration of Helsinki. Ethical Principles for Medical research Involving Human Subjects (different versions)

Documents with recommendation character

- The draft “Implementation of Good Clinical Practice Software” by JM Lauritsen, University of Southern Denmark (02/2007)
- The policy document of the German Coordinating Centres for Clinical Trials networks (October 23rd 2001, updated December 20th 2007)
- Good Clinical Data Management Practices, Version 4, SCDM, October 2005.
- CDISC standards (http://www.cdisc.org; last visited 11 August 2008)

- CONSORT – statement (Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001; 357(9263):1191-1194)
- NIHR Information Systems Enterprise Architecture, Document Number: NIHR 4.2EA001, Vers 1.0 (01/10/2007)
- MedDRA Term Selection: Points to consider (www.meddramsso.com)

Appendix III: Scenarios

Four main scenarios for data management in clinical trials

In the following paragraphs, four scenarios for data management are defined, ranging from full paper-based data management to full remote data capture with eSource. In many cases, a number of scenarios may be necessary for the same trial. Balancing advantages and disadvantages, scenario 4 seems to be a preferable option for trials related to ECRIN.

Scenario 1: fully paper-based approach

This is the classical scenario for clinical data management, in which data are collected on paper until the last step where data are entered into the statistical analysis system.

Data collection

The collection of patient data is done by the investigator at each site (Case Report Forms) or by the patient himself by completing paper questionnaires pCRF, e.g. patient diary. Any change or correction to a CRF should be dated, initialled, and explained (if necessary) and should not obscure the original entries. No automatic control of correctness of the input or of the quality of the data exists at this stage. Data collection with pCRFs is very convenient; it can be done at any time, at any place, independent from hardware or software. On the other hand, pCRFs easily reach a size of several hundred pages and take up room at the investigator's site. The completed pCRFs are collected at each site and sent to the clinical trial centre.

Data management/query management

Practically no data management exists for paper-based data, but there is a form of document management. Incoming pCRFs are checked at the clinical trial centre for completeness, legibility and obvious mistakes, marked or stamped and occasionally provided with a bar code. Queries are created at the clinical trial centre with a paper form sent to the site and the investigator has to complete the paper query and send it back to the trial centre, retaining a copy at the site filed with the pCRF. No clinical trials database exists which contains the complete set of trial data at this stage.

Monitoring

The monitor has to travel to the site and check pCRFs against patient records, examine the queries and other essential clinical trial documents.

Adverse event management and reporting

Adverse event management and reporting is done at the clinical trial centre by paper notifications. The investigator has either to complete an adverse event page of the

pCRF or a special Adverse Event form. Adverse event management has to be done by arranging, reviewing and examining the corresponding paper-based forms.

Data analysis/reporting

Only at this stage data are entered into a computer system at the clinical trial centre, possibly by using scanning and Optical Character Recognition (OCR). Minimal requirements for this procedure have to be set up. Data are imported into statistical software, which is used for analysis and reporting. Quality checks and analysis can be done easily once data are inside a database.

Scenario 2: paper-based data collection at the site, data management by CDMS at trial centre

This is perhaps the most commonly known and used case, where patient data are collected on paper forms at different sites, to be entered afterwards into the system at the central clinical trial centre.

Data collection:

The collection of patient data is done by the investigator at the site or by the patient (e.g. patient diary) by completing paper-based Case Report Forms (pCRF). As in scenario 1, any change or correction to a CRF should be dated, initialled, and explained (if necessary) and should not obscure the original entries. There is no automatic control of correctness of the input or the quality of the data at this stage. The pCRFs are collected at each site and sent continuously to the clinical trial centre.

Data management/query management

At the clinical trial centre data from the pCRFs are entered into a Clinical Data Management System (CDMS). This can be done through scanning and Optical Character Recognition, but most often it is done by Double Date Entry (DDE). An alternative would be single data entry with controls. Because data entry is performed centrally, it can be done by personnel experienced with the CDMS and with special features of data entry. During DDE pCRFs are checked for completeness and errors; the quality of data is increased through data validation routines. In case of implausible or incorrect data, electronic queries are generated, printed and sent to the investigator who has to complete the query and send it back to the trial centre. In addition, paper-based queries may be created and printed. Another option would be to use statistical analysis software for generation of queries; this may reduce plausibility checks during data entry. In the trial centre a clinical trials database exists with the complete set of data including administrative and audit trail information.

Monitoring

The monitor has to travel to the site and check pCRF against patient records, or other essential documents, and assist the query resolution process.

Adverse event management and reporting

Adverse event management and reporting is done by paper notifications, either by completion by the investigator of an adverse event page of the pCRF or by completing a special paper form. Adverse event management can be done in part at the clinical trial centre by using the CDMS or a specific software tool, after entering adverse event information into the system.

Data analysis/reporting

Simple forms of analysis and reporting (e.g. interim reports) can be done at the clinical trial centre with the CDMS. CDMS exports the clinical trials database in a format that can easily be imported by a statistics programme for further analysis and reporting.

Scenario 3: paper based data collection and remote data entry into CDMS at the site

In this case, data entry at the site consists of two steps. In many clinical situations the initial collection of clinical data on pCRFs may be the most flexible and convenient way. There is no necessity to send pCRFs and to work with paper queries.

Data collection

The collection of patient data is done by the investigator at the site or by the patient (e.g. patient diary) by completing paper Case Report Forms (pCRFs) as a first step. As in scenarios 1 and 2, any change or correction to a CRF should be dated, initialled, and explained (if necessary) and should not obscure the original entries. In a second step, the pCRFs are entered into the CDMS at the site using electronic Case Report Forms (eCRFs) by remote data entry by the investigator or an assistant. Therefore, pCRFs do not have to be sent to the clinical trial centre. The data are nearly immediately available at the central trial database located at the clinical trials centre. This database comprises the complete trial information including administrative and audit trail information. The quality of data is checked at the site by using data validation routines during data entry.

Data management/query management

Data management and query management can be done continuously during the conduct of the trial, because the database is continuously updated with new data from the sites. In case of implausible or incorrect data, electronic queries are generated automatically at the site and can be answered by the investigator online. In addition, queries generated at the clinical trial centre can be answered by the investigator directly online. In the trial centre, a clinical trials database exists with the complete set of data including administrative and audit trail information.

Monitoring

The monitor has to travel to the site and check eCRF against pCRF, pCRF against patient records or other essential documents.

Adverse event management and reporting

Adverse event management and reporting is done by using eCRFs, or by completing a special paper form. Adverse event management can be done in part by using the CDMS or a specific software tool.

Data analysis/reporting

Certain types of analysis and reporting (e.g. interim reports) can be done at the clinical trial centre with the CDMS (or a statistical analysis programme). CDMS exports the clinical trials database in a format that can easily be imported by a statistics programme for further analysis and reporting.

Scenario 4: fully remote data capture (eSource)

This scenario describes the novel case of fully electronic data collection and management.

Data collection:

The collection of patient data is done by the investigator at the site or by the patient (e.g. patient diary) by completing an online questionnaire, the electronic Case Report Forms (eCRF). Because no paper CRF exists, this scenario stands for eSource, which means that the source data exist electronically. The quality of data is increased by using data validation during remote data entry at the site. The investigator can react immediately on problematic data during the data entry step. Because no paper record exists, data are immediately available in the central clinical trials database at the trial centre. This database comprises the complete trial information including administrative and audit trail information. The audit trail exists fully electronically.

Data management/query management

Data management and query management can be done continuously with up-to-date data at the clinical trial centre because the database is continuously updated with new data from the sites. There is no delay due to the creation, management and sending of paper CRFs. In case of implausible and incorrect data, electronic queries are generated automatically and can be answered immediately by the investigator online.

Monitoring

Because no paper CRF exists, monitoring can be done to a large part by examining the database (remote monitoring). For the comparison of patient records with eCRFs, the monitor has to travel to the site.

Adverse event management and reporting

Adverse event management and reporting is done to a large degree online, either by completing an eCRF page for adverse events or by completing a special form.

Adverse event management can be supported by the clinical trial centre continuously during the conduct of the trial using updated information. SAR/SUSAR notifications are often still sent by fax and have to be entered into the system at the clinical trial centre, but may also be set to run automatically from the database. Safety management may be supported by the CDMS or a specific software tool.

Data analysis/reporting

Data analysis can be done continuously at the clinical trial centre. Simple types of analysis and reporting (e.g. interim reports) can be done with the CDMS (or a statistical analysis programme). CDMS exports the clinical trials database in a format that can be imported by a statistics programme for further analysis and reporting.

Scen-ario	Main features	Advantages	Disadvantages
1*	Paper data collection; Central electronic data entry limited to the part of the data that will be analysed.	No local hardware constraint; Professional & centralised double data entry; Entry limited to the analysed data; Reduced cost for small studies; No technical training necessary for investigator.	Heavy paper exchanges; Heavy paper archives; No automatic data control at the source; Manual data control at the centre; Automatic data control limited to the data entered in the CDMS; No complete database (e.g. queries not included); Delay between data collection and data entry, a problem for interim reports**; Danger of lack of homogeneity in data controls and corrections; Danger of global misunder-standing of a question of the CRF or protocol elements that is not detected and corrected during study; Danger that queries may not be resolved due to relay period.
2*	Paper data collection; Electronic data entry at the central DM site.	Professional & centralised double data entry; Automatic data control for the whole DB; Homogeneity in controls and corrections; Cost optimal for long term studies (> 3 years); Possibility of online reports shared between coordinator, sponsor.	Heavy paper exchanges; Heavy paper archives; No automatic data control at the source; Late detection of AE possible on the (complete) data base; Delay between data collection and data entry, a problem for interim reports.**
3	Paper data collection; Electronic data entry at the investigator site.	Decentralised data entry close to clinical expertise; Most queries are at the time of data entry; Light work for data centre; Homogeneity in controls and corrections; Cost optimal for 1-3 year duration studies; eCRF can be used for communication (newsletter, FAQ, protocol & other online documentation); Online reporting can improve recruitment.	Single data entry; Heavy paper archives; Heavy workload for investigator team (Paper data collection + electronic data entry) and for monitoring team (comparison of paper CRF against eCRF); Late detection of AE possible on the (complete) database; No automatic data control at the source; Delay between last follow-up of last patient and publication of results;Some costs due to hotline and web site maintenance.

4	No paper data collection; Continuous and direct electronic data entry at the investigator site.	Decentralised data entry close to clinical expertise; Immediate & automatic data control at the source; Light on-site monitoring; No paper archive; Ideal for continuous monitoring and / or analysis; Detection of AE possible on the complete data base; Electronic management of AE; Automatic branching and selection of question can speed up data-entry; Reduced delay between last follow-up of last patient and publication of results. eCRF can be used for communication (newsletter, FAQ, protocol & other online documentation); Online reporting can improve recruitment; eCRF can be used for communication (newsletter, FAQ, protocol & other online documentation); Easy revision of eCRF centrally if necessary (no reprint, no reinstallation for web-based applications).	Hardware and network constraints; Single data entry; No paper backup (no data when electronic system is not available, or destroyed); Inconvenience of computer or PDA screen interface during outpatient visit; Requires specific regulation on electronic source documents, not in place in all countries; Extended costs due to hotline and web-site maintenance.

Table: Advantages and disadvantages of the different scenarios

* Scenario 1 and 2 differ only by the limitation of data entry to a part of the trial data, with some delay between data collection and data entry. These are only two aspects of the same scenario.

** Delay between data collection and data entry, a problem for interim reports: scenario 4, through direct electronic data entry, reduces best this delay. The other 3 scenarios can reduce it by specific procedures, involving continuous transmission of pCRF from investigation centre for scenarios 1 and 2, or continuous data entry for scenario 3, and continuous management/entry in data management centre.

Appendix IV: Section data management of Trial Master File

Filing of all study-relevant documents compiled and received by Data Management including completed CRFs

Study documents	Comments	Essential	Optional
Lists for study documentation follow-up		☐	☐
Documentation on treatment allocation and decoding (if required) and trial subject registration including correspondence • Randomization forms • Randomization files • Randomization failures • Randomization User Guide (e.g. IVR script) • Trial subject registration (in case of non-randomized studies).	This information can also be classified in the patients files (completed CRFs).	☐	☐
Data Management Manuals/Working Instructions • Data Management Plan • Data Handling report.		☐	☐
Annotated CRFs		☐	☐
Plan for handling variables		☐	☐
Documentation of study set-up	Including modifications to versions (e.g. software release changes) in software, study-specific modifications to software/modules include validations etc.	☐	☐
Validations/revalidations • Database files • Data input masks • Programs.	All documents on the validation/revalidation including test subjects, CRF pages, notes on EDP location, separate according to data bank files, data input masks and programs.	☐	☐
Documentation data-base set-up/modification according to validation/revalidation		☐	☐

Data quality check	100% check of essential variables (e.g. target variables, header data), spot checks of remaining variables.	☐	☐
Data entry SOP or User Guide (CDMS or eCRF)		☐	☐
CRF handover protocol	From monitor to data manager.	☐	☐
Completed CRFs	Separate CRF folder.	☐	☐
Data cleaning and corrections • Queries • Medical review and corresponding files.	Queries may be filed with respective CRFs.	☐	☐
Other checks/programming	Also important external data.	☐	☐
Coding of texts (AEs, SAEs, SUSARs, medication, etc.) • Study-specific working instructions • Documentation of coding, including problems.		☐	☐
Status reports • Recruitment • CRF overview • Completeness of documentation.		☐	☐
Documentation on closing and opening of data-base		☐	☐
Database transfer to sponsor/cooperation partner		☐	☐
Documentation on outstanding/unsolved problems		☐	☐
Miscellaneous		☐	☐