DATA COLLECTION AND DATA MANAGEMENT

Pravich Tanyasittisuntorn, MD

What are basic elements of Data Quality & Data Integrity

Basic Elements of Data Quality

- **ALCOA**: basic elements of data quality
  1. **A**ccurate **ถูกต้อง**
  2. **L**egible **อ่านออก** (human readable)
  3. **C**omplete **ครบสมบูรณ์** and **C**ontemporaneous **บันทึกข้อมูลครบถ้วนทันทีที่ได้รับ** (recorded at the time the activity occurs)
  4. **O**riginal **ข้อมูลจากแหล่งกำเนิด** (original source)
  5. **A**ttributable to the person who generated the data **สามารถระบุเจ้าของข้อมูล**

Basic Elements of Data Integrity

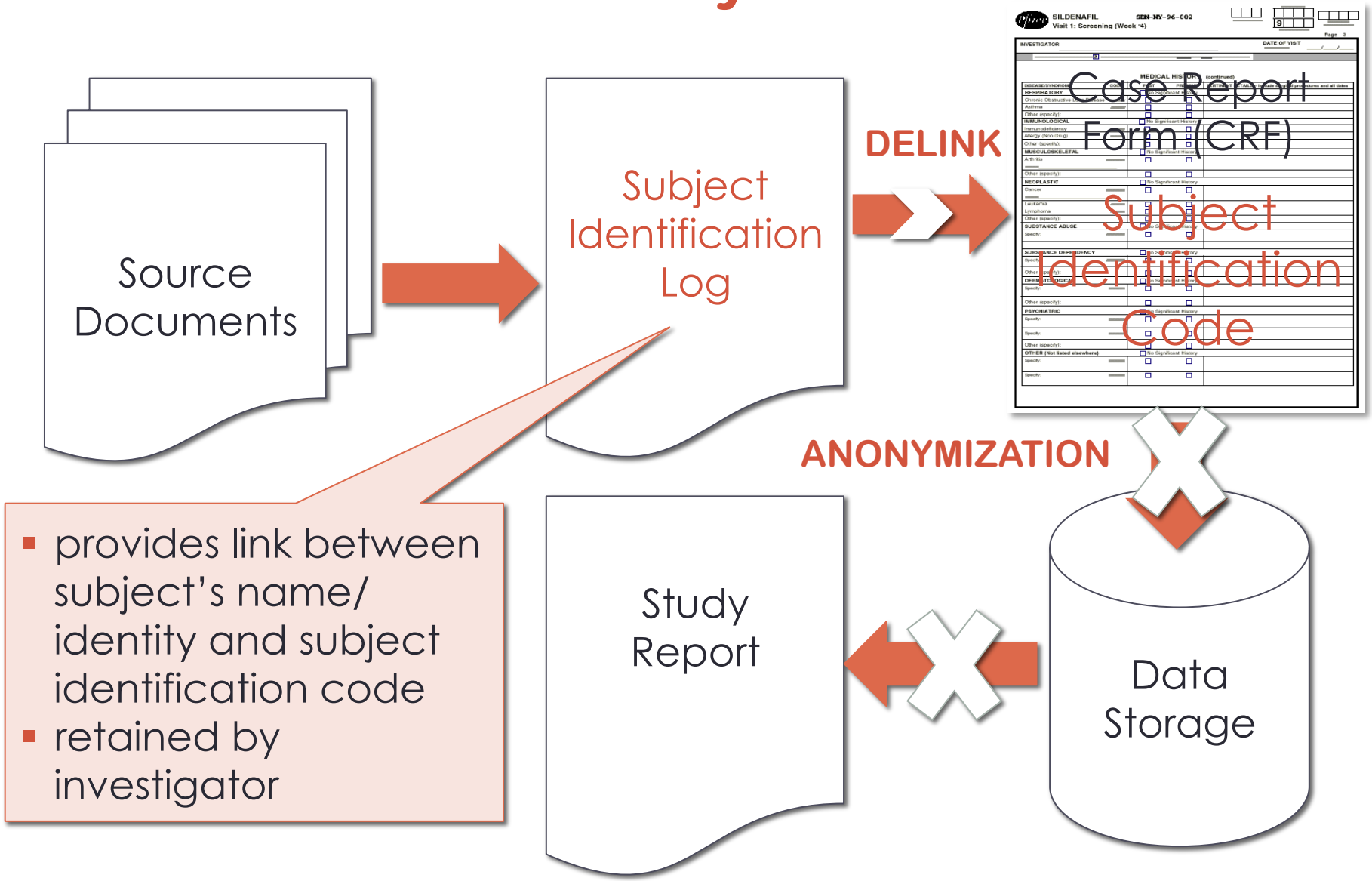
- Data integrity (trustworthiness): the degree to which a collection of data is
  1. Credible **เชื่อที่ได้** (worthy of belief and confidence, supported by known facts)
  2. Internally consistent **เที่ยงตรง** (non-contradictory, consistent terms)
  3. Verifiable **ตรวจสอบยืนยันได้**

Ethical consideration in data management

Ethical Obligations

- Protect confidentiality and privacy of subjects
  1. IRB/IEC protocol approval
  2. Data access control: only authorized persons can access (**ผู้ที่จะดูข้อมูลส่วนตัว ต้องเป็นผู้ที่อาสาสมัครอนุญาตเท่านั้น**)
  3. Collecting minimal identifiable information (**เก็บข้อมูลที่สามารถระบุตัวตนอาสาสมัครให้น้อยที่สุด**)
  4. Establishing confidentiality protection
     - delink or unlink (anonymize) data

Confidentiality Protection

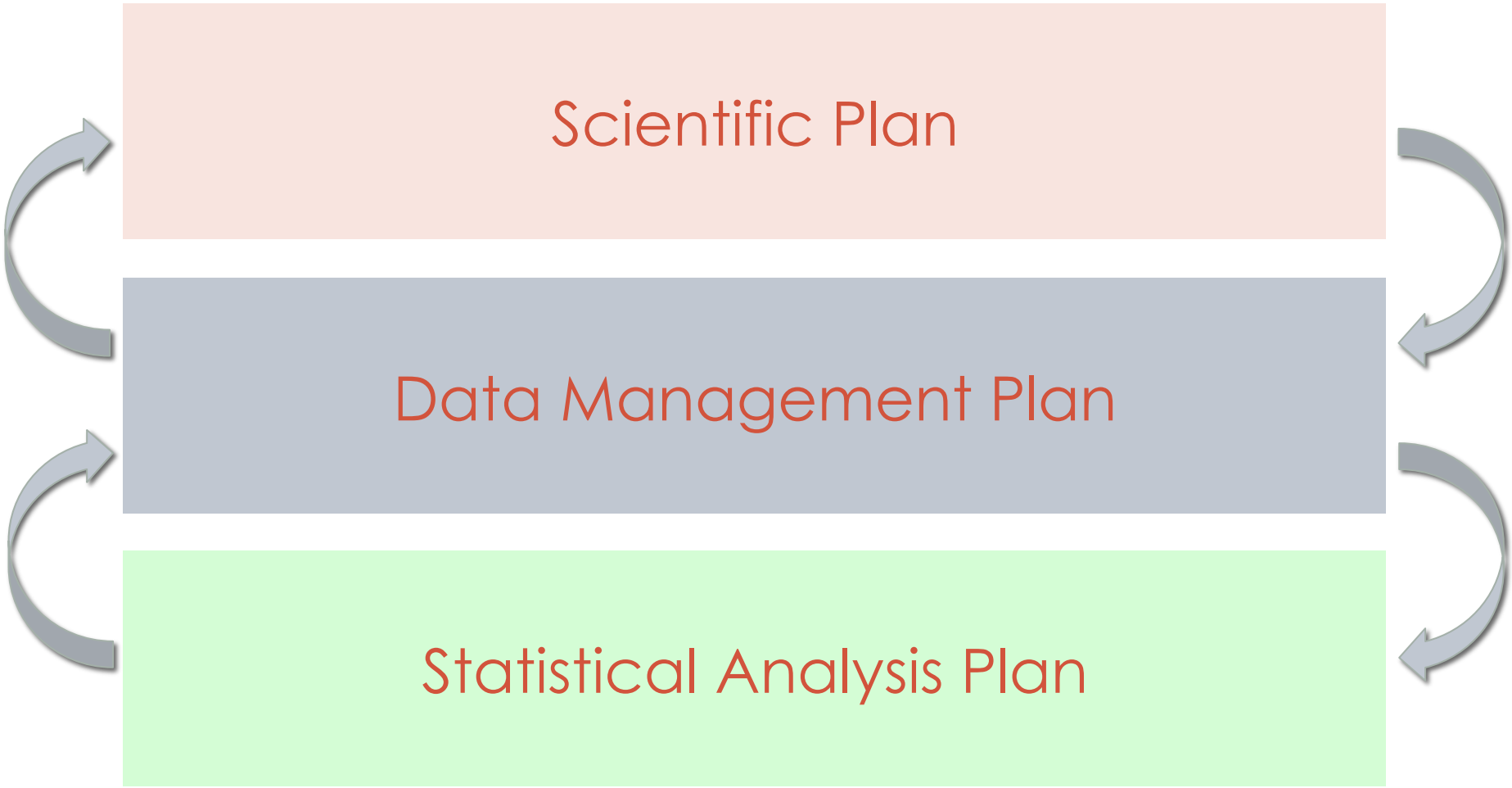


Confidentiality Protection

- Subject Identification Code
  - assigned by the investigator to each trial subject to protect the subject's identity
  - used instead of subject's name when the investigator reports adverse events and/or other trial related data.

(Ref. ICH 1.58)

Overview of data management


Scientific Plan
Data Management Plan
Statistical Analysis Plan

Scientific Plan

- Study Protocol (study proposal):
  a. Study objectives
  b. Study design, study population, study methods
  c. Determine data to be collected (data items or variables)
     - Specify the data to be collected:
       - Inclusion & exclusion criteria and basic characteristics
       - Outcomes: efficacy (i.e. clinical, lab, diagnostic, PRO) and safety
       - Associated factors (factors affecting outcomes)
     - Help to identify the characteristics of data which is required for developing CRF, database structure and statistical analysis plan:
       - numeric (continuous, discrete)
       - categorical (ordinal, nominal)

Data Management Plan

1. Data collection tool:
   - CRF (case report form)
   - Questionnaire for survey study
2. Source document plan:
   - Source documents and flow
   - Develop worksheet, if required
3. Data storage:
   - Database structure and database application

Data Management Plan

4. Data coding
   - Data dictionary
   - Annotated CRF
   - Textual data e.g. disease, drug, adverse event terms
5. Data flow and CRF tracking
6. Data entry:
   - Manual
   - Scanning
   - EDC (electronic data capture)

Data Management Plan

7. Data cleaning / data validation / edit checks
   - identify
     - missing page
     - obvious data error, data inconsistency, non-meaningful data, omitted / missing data, additional hand written data
     - protocol deviation
   - time-of-entry and back-end edit checks
   - review data in CRF point-by-point and page-by-page

Data Management Plan

7. Data cleaning / data validation / edit checks
   - Activities
     - review CRF pages
     - manually review data
     - identify discrepancies through computer generated checks / validation
     - conduct logical and statistical review checks

Data Management Plan

8. Managing discrepancies
   - Query generation and resolution
   - Protocol deviation
   - Data integration
9. Managing lab data and safety data
   - Lab test unit
   - Lab normal range
   - Medical review for abnormal test
   - Data transfer, if lab database is separated

Data Management Plan

10. Managing safety data
    - AE terms
    - Medical review for safety data completion and causality assessment
    - SAE reconciliation
    - Continual review
11. Database locked (at interim analysis and study completion)
12. Data analysis and data quality review

Statistical Analysis Plan (SAP)

a. Study objectives and variables
b. Hypothesis to be tested
c. Primary end points & secondary end points
d. Definition of analysis populations:
   - intention-to-treat, per-protocol, and safety population
e. Statistical methods to be used
   - be specific: what covariates, compare what treatment groups, in what FU visits)
f. Subgroup analysis
g. Data handling rules: missing data
h. Data presentation:
   - Tables, figures and line listings (TFLs)

Investigator's responsibilities

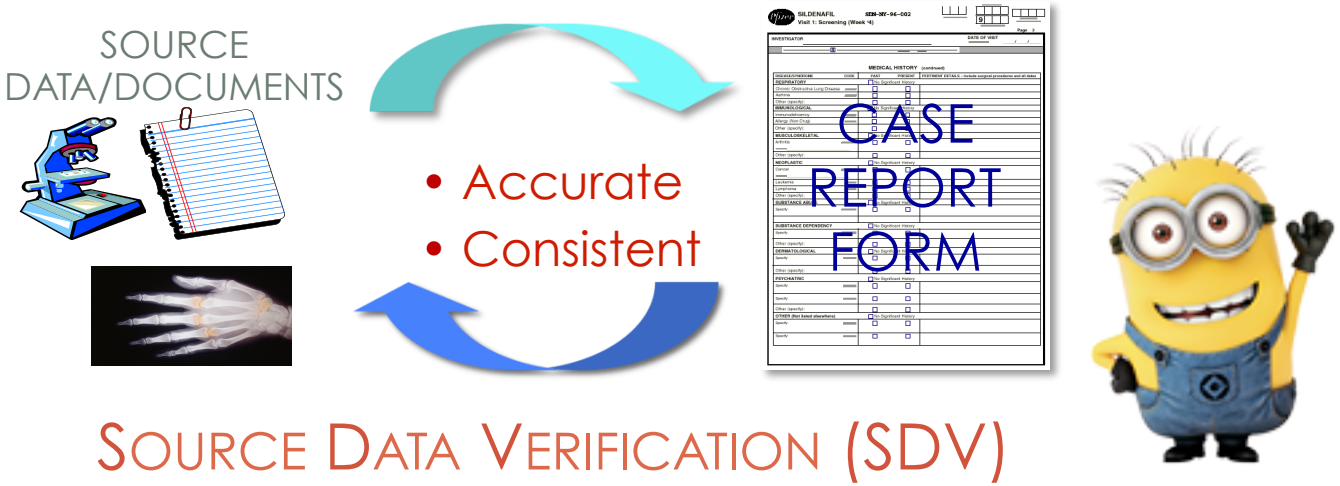
Investigator's responsibilities

- Should ensure the accuracy, completeness, legibility, and timeliness of the data reported to the sponsor in the CRFs and in all required reports

(Ref ICH GCP 4.9.1 – 4.9.6)

Investigator's responsibilities

- Data reported on the CRF, which are derived from source documents,
  - should be consistent with the source documents or
  - the discrepancies should be explained



SOURCE DATA VERIFICATION (SDV)

(Ref ICH GCP 4.9.1 – 4.9.6)

Investigator's responsibilities

- Any change or correction to the CRF should be date, initialed, and explained and should not obscure the original entry
  - Data correction:

Visit date ~~31 MAR 2012~~

Visit date ~~31 MAR 2012~~

Visit date

Visit date ~~31 MAR 2015~~ 02 APR 2015 PT 19/APR/15

(Ref ICH GCP 4.9.1 – 4.9.6)

Investigator's responsibilities

- Document and endorse any changes or corrections in CRFs made by sponsor
  - The investigator should retain records of the changes and corrections

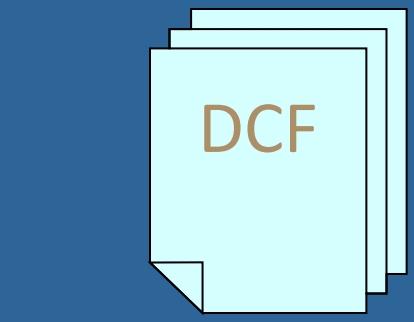
(Ref ICH GCP 4.9.1 – 4.9.6)

If it is not written,

it did not happen!!

Data collection & Data management process

Data Flow and Data Quality Control

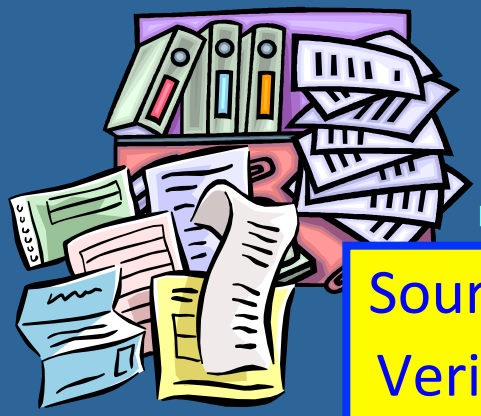


Investigator

Data clarification

Data Clarification Form

Self-evidence correction



Source Data Verification

CRF

Data validation

Data entry

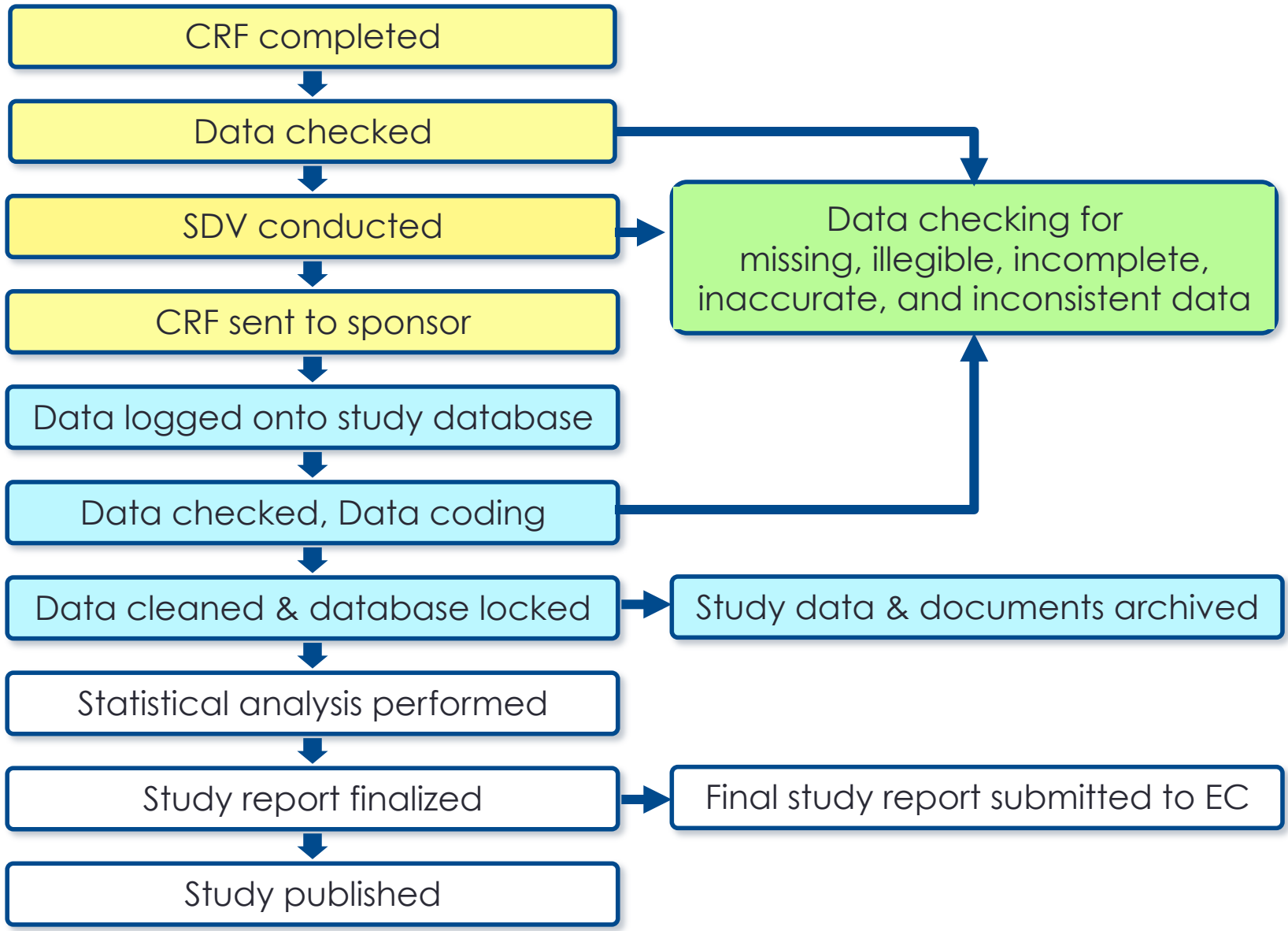
Source document

Study nurse/
Site coordinator

Investigator review & sign

Data management

Data Flow and Data Quality Control



CRF Design and Instruction for CRF completion

Source Documents

Guidance for data entry and processing

Data Entry Guidance

1. Use tracking systems for e.g., enrollment, CRF, data query
2. Enter data as soon as possible after it is received
3. Process data cleaning throughout the study in order to have early resolution of data queries
4. Identify missing CRF pages
5. Identify lab data which is out of the normal range or missing
6. Code medical or adverse event terms frequently

Key things to remember

- Know protocol well: scheduling visit, activities and data will be collected
- Review CRF and its completion guidance before study started
- Record data in sources document: “If it’s not written, it did not happen”
- Complete CRF in a timely manner: accuracy, completeness, legibility, consistency, and verifiability of data are crucial