

Datenbereinigung

Dokument Nummer	CIO-DM-009-00	Author	G Antony
Versions Nummer	3.2	Reviewer	E Sauer
Vorige Version	3.1	Review	23-Jul-2024
Tritt in Kraft am	01-Aug-2024	Status	Final

Inhaltsverzeichnis

1. Allgemeine Informationen
2. Verantwortlichkeiten
3. Definitionen und Abkürzungen
4. Methode
5. Anhänge und Formulare zur Vervollständigung
6. Verweise auf andere SOPs
7. Genehmigung und Verbreitung
8. Referenzen

1. Allgemeine Informationen

1.1. Ziel und Umsetzung

- Das Ziel dieses Vorgehens ist es, alle Hauptaspekte für die Datenvalidierung zu beschreiben. Es ist nicht das Ziele eine Arbeitsmethode Schritt für Schritt zu spezifizieren sondern vielmehr einen Rahmen zu liefern, in dem eine Arbeitsanweisung entwickelt werden kann.
- Datenbereinigung liefert Informationen zu Einschränkungen der Daten basierend auf spezifischen Qualitätskriterien, -zielen und -indikatoren. Basierend auf diesen können Diskrepanzen in den Studiendaten identifiziert und ausgebessert werden.
- Diese SOP ist für eine manuelle Prüfung der Daten und wird es ermöglichen Diskrepanzen zu finden, Queries zu stellen und zu korrigieren oder zu validieren.

1.2. Gesetzgebung und Standards

- Für klinische Studien basieren die Minimalstandards auf ICH-GCP (<http://www.ich.org/>)
 - Good Clinical Practice (GCP) ist ein internationaler, ethischer und wissenschaftlicher Qualitätsstandard zur Entwicklung, Durchführung, Dokumentierung und Berichterstattung von Studien, welche die Teilnahme von Menschen als Testsubjekte umfassen.
 - Einhaltung dieses Standards versichert öffentlich, dass die Rechte, die Sicherheit und das Wohlergehen der Studienteilnehmer geschützt werden und die Daten der klinischen Studie glaubwürdig sind.
- Für elektronische Datenerfassung definiert 21 CFR Teil 11 die Kriterien, unter welchen elektronische Aufnahmen und Signaturen vertrauenswürdig, zuverlässig und als ebenbürtig zu Papieraufzeichnungen angesehen werden.
- Für nichtklinische Studien werden die Minimalstandards von Studie zu Studie, risikenbasiert, mit Zustimmung des Vorsitzenden der die Studie durchführenden Organisation, des Projektleiters der Studie und des verantwortlichen Daten-Managers bei CIO entschieden, während das allgemeine Ziel der Datensicherung vom Quelldokument bis zur Datenbanksperrung eingehalten wird.
- Falls örtliche Gesetzgebung zusätzliche Standards des DM fordert, müssen diese übernommen werden.

2. Verantwortlichkeiten

Rolle	Verantwortlichkeit
Leitung von CIO Marburg	<ul style="list-style-type: none"> • Bewilligt Datenvalidierung
Daten-Manager	<ul style="list-style-type: none"> • Führt Aufgaben aus, die in dieser SOP genannt werden • Erstellt einen Datenvalidierungsplan
Systementwickler	<ul style="list-style-type: none"> • Baut automatische Datenvalidierung in das Daten Management System ein
Projektleiter oder Vertretung	<ul style="list-style-type: none"> • Bewilligt diese SOP und ihre Veränderungen, bevor sie in der Institution genutzt wird • Autorisiert selbstverständliche Korrekturen
Dateneingabepersonal oder Prüfleiter	<ul style="list-style-type: none"> • Löst Diskrepanzen auf • Beantwortet Queries
Prüfer	<ul style="list-style-type: none"> • Prüft die Dateneingabe insbesondere auf Diskrepanzen • Stellt Queries, die sich auf die Studiendaten beziehen
Studienstatistiker	<ul style="list-style-type: none"> • Hilft, einen Daten Management Plan zu erstellen

3. Definitionen und Abkürzungen

Definitionen

- Diskrepanzen: alle Inkonsistenzen in der klinischen Studie, die Untersuchung fordern
- Datensperre: Änderung des Status der Daten, sodass diese nicht weiter geändert werden können
- Query: Nachfrage eines Prüfers zur Dateneingabe

Abkürzungen

CDMS	Clinical Data Management System (Klinisches Daten Management System)
CRF	Case Report Form (Prüfbogen)
DCF	Data Clarification Form (Formular zur Datenklärung)
EDC	Electronic Data Capture (Elektronische Datenerfassung)
PID	Participant ID (Teilnehmer ID)
SEC	Self-Evident Corrections (Selbstverständliche Korrekturen)
SOP	Standard Operating Procedure (Standardarbeitsanweisung)

4. Methode

4.1. Diskrepanzen identifizieren

- Diskrepanzen identifizieren, die aus verschiedenen Quellen stammen können, wie z.B.:
 - Editchecks, die in das EDC System eingebaut sind
 - Editchecks von Daten außerhalb des Dateneingabesystems (z.B. Auflistungen von Statistikern)
 - Manuelle Prüfung
- Automatische Editchecks
 - Erstellung einer Liste (z.B. Excel Tabelle) aller Editchecks, die automatisch programmiert sind (in dem EDC System oder außerhalb mit einem Statistikprogramm). Sie können das Data Dictionary als Vorlage nutzen, das während der eCRF Erstellung genutzt wurde (s. SOP Datenbank und eCRF Entwicklung).
 - Editchecks sollen nicht die Dateneingabe verhindern sondern auf Inkonsistenzen hinweisen, da dies zu fehlenden Daten führen könnte

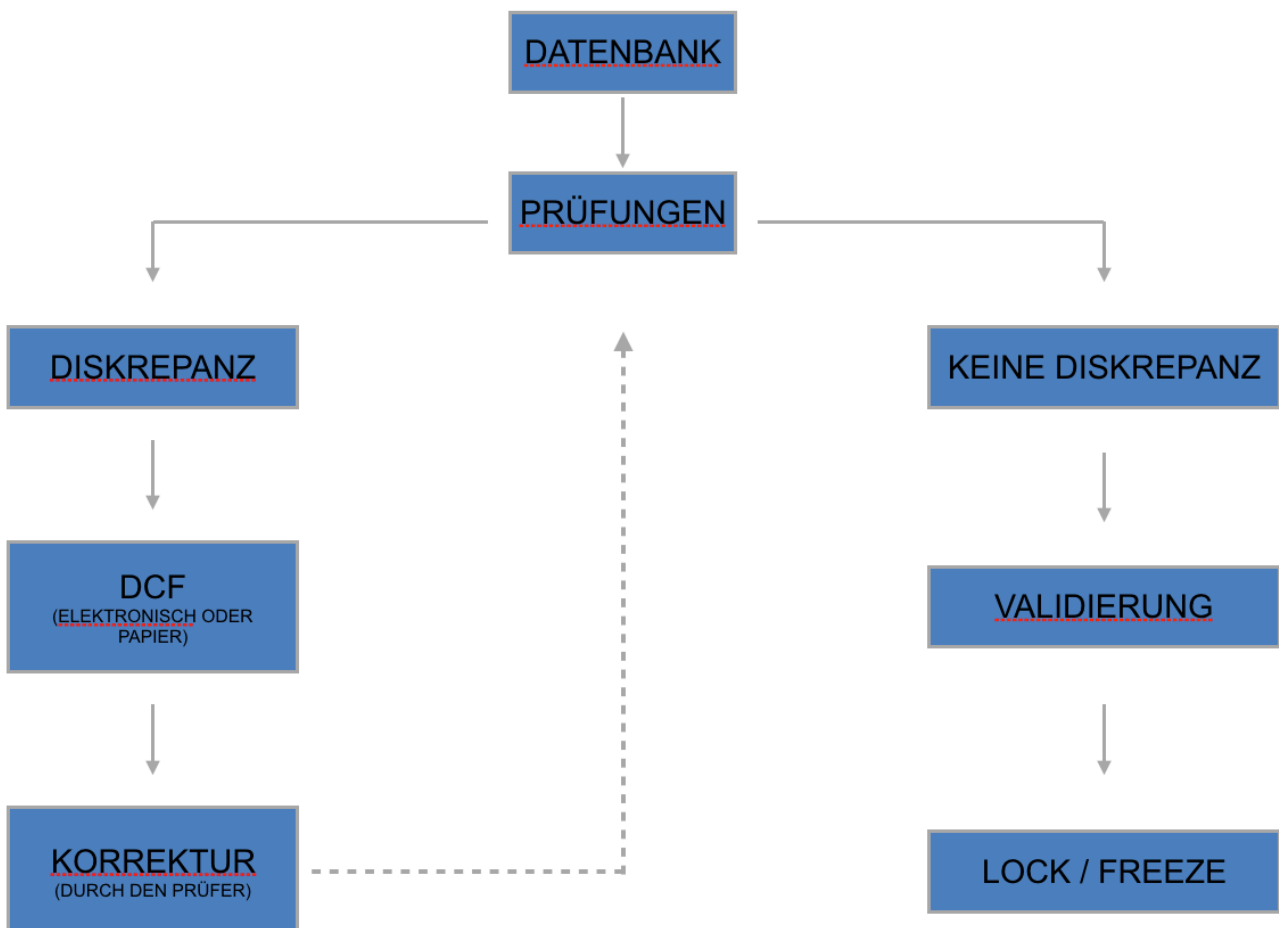
- Die Arten der Editchecks in folgender Liste sind einige Beispiele von vielen anderen:
 - Fehlende Werte
 - Fehlende Patientenaufzeichnungen
 - Bereichsprüfung (z.B. Körpergröße zwischen 150 und 200 cm)
 - Logische Fehler (z.B. Laborwerte sind als nicht fertig bestätigt während Laborwerte eingegeben werden)
 - Gegenprobe zwischen Modulen und Visiten (z.B. Reihenfolge der Visitedaten)
 - Protokollmissachtungen (z.B. Inkonsistenz in Einschlusskriterien)
- Manuelle Prüfungen
 - Erstellung einer Liste (z.B. Excel Tabelle) aller manuellen Prüfungen, die an den Studiendaten durchgeführt werden sollen. Sie können das Data Dictionary als Vorlage nutzen, das während der eCRF Erstellung genutzt wurde (s. SOP Datenbank und eCRF Entwicklung). Die Arten der Prüfungen in dieser Liste sind einige Beispiele von vielen anderen.
 - Freier Text
 - Gegenprüfung von concomitanten Medikationstabelle und einer Tabelle unerwünschter Ereignisse
 - SAE Abgleich
 - Prüfung der Visitedatenreihenfolge

4.2. Diskrepanzverwaltung

- Diskrepanzen, die als SEC klassifiziert sind, identifizieren und auflösen. Beispiele sind:
 - Fehlende Werte können in manchen Fällen übersprungen werden (z.B. Frauenfragen für Männer)
 - „Wurde Laborprobe entnommen?“ ist als Nein markiert oder leer obwohl Laborwerte vorhanden sind.
 - Korrektur falls Jahreswerte falsch aber eindeutig ein Schreibfehler sind.
- Falls eine Query gestellt werden muss ist es wichtig, dass die Frage den Prüfer nicht zu einer bestimmten Antwort leitet
 - Z.B.: Nicht „Ist dieser Wert 5?“ sondern „Bitte diesen Wert auf dem CRF prüfen und bestätigen.“
- Konsistenz zwischen Queries erhalten. Falls möglich Vorlagefragen erstellen, die für ähnliche Fälle genutzt werden können.
- Falls eine Antwort nicht der Query genügt, ist diese umzuformulieren und zu verdeutlichen, damit am Standort keine Frustrationen ausgelöst werden.
- Vermeiden Sie wütende Queries zu schreiben, halten Sie die Wortwahl genau und professionell.

- Queries verfolgen
 - Eingebautes Diskrepanzverwaltungssystem nutzen, dass es ermöglicht Queries in die Eingabemaske zu schreiben und daher automatisch zu verwalten. Ändern des Status einer Frage von „gestellter“ Diskrepanz zu „beantworteter“ Diskrepanz bis der Status „geschlossen“ ist.
 - Falls ein Datenverwaltungssystem genutzt wird, das keine Diskrepanzen verwaltet, müssen Diskrepanzen parallel mit DCFs eingesendet werden. (s. DCF Vorlage)
 - Änderungen der Daten anhand des DCF in der Datenbank vornehmen
 - DCF zusammen mit dem Papier CRF in der Patientenakte aufbewahren.

Alle Daten, die durch manuelle Prüfung oder Editchecks bereinigt wurden und als geschlossen bewertet sind können gesperrt werden, falls das Datenverwaltungssystem es erlaubt. Allgemeiner Fluss s. Abbildung:



4.3. Datenvalidierungsplan

- Erstellung eines Datenvalidierungsplans, in dem der Output von 4.1. und 4.2. verbunden wird, und gegebenenfalls eine Erweiterung der Zeitpläne erfolgt

5. Anhänge und Formulare zur Vervollständigung

Anhang 1: DCF (Formular zur Datenklärung)

6. Verweise auf andere SOPs

Dieses SOP konzentriert sich auf Datenbereinigung und sollte gemeinsam gelesen werden mit:

- CIO-DM-007-00-SOP-Daten-Management-Plan
- CIO-DM-004-00-SOP-System-Validierung
- CIO-DM-010-00-SOP-Datenkodierung
- CIO-DM-003-00-SOP-Datenbank-und-eCRF-Entwicklung
- CIO-DM-012-00-SOP-Daten-Nachverfolgung

7. Genehmigung und Verbreitung

	Name und Funktion
Initiiert von:	Gisela Antony, Leiterin CIO Marburg
Überarbeitet von:	Edda Sauer, CIO Marburg Personal
Manuelle Verbreitung:	Nur für momentanen Druck bestimmt

8. Referenzen

Practical Guide to Clinical Data Management (Susanne Prokscha – third edition)