

SDTM – Erfahrungen beim Mappen mit Testdaten

Konstanze Morgenroth & Stefanie Sturm

Accovion GmbH

18. Februar 2014



Inhalt

- ✓ Warum?
- ✓ Eigene Testdaten
- ✓ Testdaten vom Kunden
- ✓ Fazit

Mappen mit Testdaten – Warum?

✓ Timelines vom Kunden

- Genügend produktive Daten erst spät verfügbar

✓ SDTM-Daten als Grundlage für Auswertung

- Frühes Erstellen von Spezifikationen, Analysedatensätzen und Tabellenprogrammen

✓ SDTM-Daten als Grundlage für weitere Arbeiten

- Frühes Erstellen von Programmen ...
 - für die Bestimmung von Protokollverletzern
 - für Manual/Medical Review Listings
 - zur Analyse des Fortschritts einer Studie
 - als Grundlage für neue, in Planung befindliche Studien

Mappen mit „eigenen“ Testdaten



Welche Testdaten stehen zur Verfügung?

- ✓ Testdaten um Datenbank-Setup zu überprüfen
 - ✓ Testdaten um Edit Checks zu testen
 - ✓ Testdaten um externe Data Loads zu überprüfen
- 

Testdaten um Datenbank-Setup zu überprüfen

✓ Ausgangslage:

- Meist nur 1-2 Patienten

✓ Problematik:

- zu wenig Daten für sinnvolles Mappen
 - schwierig nachzuprüfen, ob Programmierung für alle Ausprägungen funktioniert
- 

Testdaten um Edit Checks zu testen

✓ Ausgangslage:

- Es werden bewusst „falsche“ Daten eingegeben
- Es werden nur die jeweiligen Seiten, aber keine kompletten CRFs eingegeben

✓ Problematik:

- hinderlich für korrektes Mappen
 - lückenhafte Datensätze beim Mappen
- 

Testdaten um externe Data Loads zu überprüfen

✓ Ausgangslage:

- Werden von externen Anbietern (z.B. Labore) geliefert
- Testdaten um Datentransfer zu prüfen und zu finalisieren
- Oftmals recht spät, wenn Mapping mit produktiven Daten der restlichen SDTM Datensätze bereits begonnen hat

✓ Problematik:

- Daten manchmal fehlerhaft und erst bei letztem Testload sinnvoll nutzbar
- Häufiger Wechsel zwischen produktiven und Testdaten
- Mapping kann nicht (final) abgeschlossen werden

Wechsel zu produktiven Daten

✓ Ausgangslage:

- Irgendwann genügend produktive Daten verfügbar

✓ Problematik:

- Andere Patientennummern, da z.B. external data noch auf Testdaten basieren
 - Site Numbers (XDUMMY vs. ABC123)
 - Uncleane Produktivdaten
- 

Mappen mit Testdaten vom Kunden



Mappen mit Testdaten vom Kunden (1)

Ausgangssituation:

- ✓ Testdaten wurden von einem dritten Partner für den Kunden erzeugt; mit diesen wurde von uns das SDTM Mapping programmiert.
- ✓ Ziel: SDTM Datensätze stehen frühzeitig zur Verfügung, damit früh mit der Erstellung von Auswertungsprogrammen begonnen werden kann.
- ✓ Prioritäten bei den Domains, die vom Kunden vorgegeben wurde

Mappen mit Testdaten vom Kunden (2)

Umsetzung:

- ✓ mehrere Testdatenlieferungen abhängig von den Prioritäten der SDTM Domains
 - ✓ Erzeugung der Testdaten erfolgte ‚randomly‘
- 

Mappen mit Testdaten vom Kunden (3)

Auswirkungen:

- ✓ Änderungen der Testdaten nach Absprache möglich, teilweise aber auch unerwartete Änderungen in neuen Datenlieferungen
- ✓ Kein logischer Zusammenhang innerhalb eines Datensatzes
 - unerwartete doppelte records
 - unerwartete Einträge in ‚abhängigen‘ Variablen (z.B. Subject completed study = Yes, aber Grund für Abbruch vorhanden)
 - ...

Mappen mit Testdaten vom Kunden (4)

- ✓ Kein logischer Zusammenhang zwischen Datensätzen
 - Unterschiedliche Patientennummern in den einzelnen Datensätzen, merge problematisch
 - Datumsangaben zu einem Patienten passen nicht zusammen
 - --DY Variablen konnten nicht mit RFSTDTC berechnet werden
 - ...

Mappen mit Testdaten vom Kunden (5)

✓ Datumsfelder

- Unerwartete partial dates
- Finales Datumsformat unklar (Trennzeichen ,/‘ oder ,-‘)

✓ Testdaten enthielten viele Daten, die nur in SUPP-- Datasets abgelegt werden können (z.B. Mehrfachankreuzung bei Tickboxen mit ,tick all that apply‘).

- Spezifizierung/Programmierung von SUPP– könnte sich am Ende als völlig überflüssig herausstellen.
- Anpassung der CT

Mappen mit Testdaten vom Kunden (6)

✓ Externe Daten (IVRS, Labor)

- Die Struktur der externen Daten war unklar, daher Erzeugung von Testdaten schwierig
- In den Testdaten für Labor passten Einheiten nicht zu den Parametern
 - Standardisierung / Anwendung von Standardmakros nicht möglich

Mappen mit Testdaten – Fazit (1)

- ✓ Viele Fehlermeldungen beim Programmieren, die immer wieder überprüft werden müssen (Fehler durch uncleanen Testdaten oder tatsächlich Programmierfehler ?)
- ✓ Rückfragen durch Programmierer für die Auswertung:
 - Wegen Errors/Warnings in den log-files der Mapping Programme wurden dort Fehler vermutet
 - SDTM Datensätze nicht für alle Variablen gefüllt (z.B. fehlende Standardisierung der Labordaten); Programmierung der Auswertung schwierig

Mappen mit Testdaten – Fazit (2)

- ✓ Testdaten konnten spätere produktive Daten nicht 1:1 widerspiegeln
 - ✓ Meist zu wenig Patienten insgesamt, oder zu wenige Ausprägungen, um alle Mappings zu testen
 - ✓ Mehr Aufwand, da die Programme immer wieder angepasst und überprüft werden mussten
 - ✓ Genaue Vorgaben für Testdaten wären notwendig gewesen
- 

✓ Konstanze Morgenroth
Senior Database Programmer

konstanze.morgenroth@accovion.com

✓ Stefanie Sturm
Principal Statistical Programmer

stefanie.sturm@accovion.com



Accovion GmbH
Helfmann-Park 10
D-65760 Eschborn

