Using OpenCDISC to Prepare for FDA Submissions

Regulatory requirements and common data issues

Sergiy Sirichenko

September 9, 2014

PINNACLE 21

P21

Disclaimer

The views and opinions presented here represent those of the speaker and should not be considered to represent advice or guidance on behalf of the Food and Drug Administration.

Topics

- Regulatory requirements
- General recommendations
- Common data issues

P21

Be aware of regulatory expectations

- ICH
- FDA Guidances
- New FDA requirements
  - CDER Common Data Standards Issues
  - Study Data Technical Conformance Guide
  - Upcoming data validation rules
- CDISC Standards

- Sponsor should comply with Regulatory Guidance and Standards
- Sponsors can outsource work, but not their responsibility for data quality
  - “trust but verify”
- High data quality can be achieved only by well-designed process
- OpenCDISC Validator is a tool to help, but process is also important

Regulatory landscape is changing

- CDISC Standards are becoming required
- FDA reviewers want CDISC
- FDA is starting an enforcement of Standards and Data Quality
- Science, technology and industry are ready
- There are no more excuses to avoid CDISC

Common challenges during submission preparations

- Studies are done by different vendors
  - Inconsistency in Standards interpretations
- CSRs are done based on legacy data
  - Risk of loosing traceability and getting different results
- Missing documentation
- “Too late to fix”

P21

Recommendations

- Design your process
- Start CDISC implementation early
  - IND, not NDA
  - CRF design
  - Map to SDTM before data collection
  - Prescriptive metadata
  - Run validation during data collection to fix issue

- Build up internal expertise
- Validate deliveries from vendors
- Ensure completeness of submission data packages in advance
- Use the most advanced science and utilize features of new versions of
  - Standards
  - Validation Rules

- Fix all identified data issues if possible
- Provide explanations for false-positive validation messages
- Provide explanations with sensitivity analysis on non-fixable issues
- Communicate with FDA on their expectations

Most common data issues

P21

#1 Metadata

- Most commonly overseen area
- Data without definition is useless
- Poor quality of metadata reduce quality of data

P21

Define.xml Origin

- Missing Origin
- Inconsistency. E.g.,
  - Origin=”CRF Page x” and populated Method
- Origin=”CRF” without reference to particular page

CRF

- Annotations as a highlighted text
- Annotations to original EDC variables

P21

Reviewers Guide

- Usage of outdated OpenCDISC versions.
  E.g., v1.3
  - It's not an excuse for low quality data
- Meaningless explanations for issues.
  E.g.,
  - "Expected result"
  - "As received from our vendor"

P21

Additional documents are welcome

- Supplemental data definition
- Complicated algorithms
- Mapping specifications
- Conversion and look-up tables

#2 Control Terminology (CT)

- Ignoring extensible CT. E.g.,
  - "Sponsor defined CT"
  - "We can add whatever we want"
- You can only add terms that are not already represented in standard CT
- If it's not CT compliant, it's not CDISC data

- Collecting data as a free text
  - Mapping problems
  - No control
- Converting standard terms into Upper Case

P21

Recommendations

- Start using CT during CRF design. E.g.,
  - Pre-specified options instead of free text
  - AEACN="DOSE CHANGED". Is it increased or decreased?
  - Any mapping introduces a risk of loosing or changing data

#3 Additional FDA Requirements

- EPOCH
- Study Days
- Baseline Flag
- AETRTEM flag
- Trial Summary and Trial Design domains
- Re-sizing variable length to actual max value

#4 Usage of outdated standards

- Example:
  - Oncology domains were introduced in SDTM IG v3.1.3
  - Sponsors try to implement them within SDTM IG v3.1.1
  - New variables are kept in SUPPQUAL
  - RELREC links SUPPQUAL records
  - Data is not reviewable and not CDISC compliant

P21

- FDA wants new variables. E.g., DTHFL, full MedDRA coding
- Problems with CT. E.g.,
  - VSLOC="LEFT ARM" instead of
  - VSLOC="ARM", VSLAT="LEFT"
- Usage of old standard is not an excuse for low quality data!

P21

#5 Programming/Mapping

- Inconsistent Standard Units
- Leading space characters. E.g.,
  - “ Glucose”
- <Line break> symbol
- Usage of Actual time instead of Planned for –TPT
- Usage of –STRF, --ENRF for Screen Failures

- Inconsistency in paired variables
  - --TESTCD/--TEST
  - QNAM/QLABEL
  - –TPT/--TPTNUM
  - VISIT/VISITNUM
- Inconsistency in paired CT. E.g.,
  - EGTESTCD=”QTC” and EGTEST=”QT Uncorrected”

P21

- Comments in SUPPQUAL domains
- Missing TS and Trial Design domains
- Incorrect order of variables and non-standard labels
- Data management info in submission data. E.g.,
  - Samples or scans tracking
- Imputation of study days for partially missing dates

P21

#6 Data Collection

- Duplicate records. E.g.,
  - LB results on the same test and time point as “Normal” and “Abnormal”
- AE Seriousness Criteria were not collected
- Missing Original Units for Lab Test Results

P21

Recommendations

- "Risk Based" monitoring
  - Safety data
  - Baseline/End Point records
  - Death info
- "Quality by design"
- Catch errors as soon as possible
- Start data validation early

P21

New check examples

- Sponsor is responsible for data quality
- OpenCDISC can help
- SDRG expect Sponsors to execute their additional validation checks
- Data quality is defined not by low number of identified issues, but by actual data content
- New checks are helpful, not a threat

- "If any of AE Seriousness Criteria variables has value 'Y', then AESER must be 'Y' " – Error

- "CMDECOD value is expected to be populated" – Warning
  - CMDECOD is Permissible variable
  - When its value is missing it's usually due to wrong type of info. E.g.,
    - 'Lung Biopsy'
    - 'Colonoscopy'

Questions

Sergiy Sirichenko
ssirichenko@pinnacle21.net