FIT FOR USE PILOT COLLABORATION OF:
CDISC, PHUSE, FDA

INDUSTRY FEEDBACK:
Domain / Data File Level Learning

PREFACE

- Learning points from this pilot are available to all. A publicly accessible wiki site for sharing the learning from this pilot is: https://wiki.cdisc.org/display/NSFFUW/Nonclinical+%28SEND%29+Fit+for+Use+Workstream+Home
- One of the main tenets of the Fit for Use Pilot was that the feedback provided back to each participant would be shared with the public broadly, via the CDISC and PhUSE organizations working processes. In addition to this presentation, the following deliverables are available at the above address:
  - INDUSTRY FEEDBACK, March 2017
  - CDISC-PhUSE Fit for Use Pilot, October 2016
  - SEND Challenges, Tabular Examples from Fit for Use, October 2016
  - Informational packages that include the actual (redacted) feedback from the FDA to individual sponsor participants, April 2017
- Learning from this pilot is being used to inform industry CDISC and PhUSE team efforts. These teams are open to all who are interested in data standards and implementation
- Learning from this pilot has been incorporated in version 3.3 of the Study Data Technical Conformance guide.

Contents

- Process used to assess feedback comments from reviewers to create the content of this presentation
- Learning points by domain
- Summary Learning Statements collected at the end of the pilot from the participants

Process Used

- Feedback from FDA reviewers to sponsors was aggregated and anonymized
- A subteam of pilot participants including sponsors, vendors, CRO, SEND subject matter experts and FDA reviewed the aggregated feedback, focusing first on the highest priority domains for improvement in usability, as determined by CDISC with FDA input: PC, PP, MA, MI
- The subteam then looked at additional domains as time allowed
- This presentation contains the learning points from the subteam assessment and full team review
- The subteam assessment included the following domains or data files:
  - MA, MI, and associated supplemental qualifiers
  - PC, PP
  - CL
  - EX
  - TX
  - Define
  - nSDRG
  - DM

Learning Points: CL Domain

- CL domain is a findings domain, therefore findings associated with injection sites are appropriate here. The location of injection site would be appropriate in EX.
- FDA is stating difficulties with inconsistent content when populating CL domain. This is a recognized issue and the FDA will continue to see these inconsistencies with V3.0. FDA should use ORRES rather than STRESC in the absence of a codelist

Learning Points: CL Domain (cont.)

- The FDA would prefer not to have extraneous operational data (e.g., mortality check, study director approval, dose site marked with indelible ink, "refer to comment", etc.) in the CL domain
- FDA reviewers should expect to see Ophthalmology (e.g., Ophthalmoscopy observations) in the CL domain
- Make sure the content is appropriate for the variable (regardless of what your LIMS systems produces) (for example, "postdose" is not a CL location (CLLOC) and is not traceable to the study report for CL)

Learning Points: PC, PP Domains

- Nominal timings are not consistently represented or are missing. The ability to visualize the data is difficult. It is a high priority of the CDISC SEND team to address these issues. A subteam of SEND is focusing on short-term and long-term solutions for PC, PP and related EX data, including timing needs, in 2017. The pilot served to remind us of the importance of these efforts and enabled us to provide suggestions to this subteam.

- The PC domain should support creation of time series graphs and automatic calculation of pharmacokinetic parameters from sets of related plasma concentrations. Three elements are necessary:
  - Nominal timings relative to the dose in numeric or ISO 8601 format
  - Grouping of each different set of time series measurements used to calculate a rekated (sic) pharmacokinetic parameter
  - Identification of the start of each time series relative to the start of exposure

  If the nominal times are provided in PCELTM, nulls should be avoided.*

* Excerpt from Study Data Technical Conformance Guide v3.3 dated March 2017, available on the FDA website, "Study Data Standards Resources", https://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm#guides

Learning Points: PC, PP Domains (cont.)

- The unit in PCSTRESU applies not only to PCSTRESC and PCSTRESN but also to PCLLOQ.
- When a measurement is identified as being above or below a limit or quantitation threshold in PCSTRESC and/or PCLLOQ, standardized units for the threshold should be provided in PCSTRESU.*
- Whenever pharmacokinetic parameters are reported as part of a study, it is expected that the PP domain will be included.

* Excerpt from Study Data Technical Conformance Guide v3.3 dated March 2017, available on the FDA website, "Study Data Standards Resources", https://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm#guides

Learning Points: MI, MA Domains

- Use RESMOD according to the definition provided. (e.g. do not put the entire ORRES finding into STRESC when it contains elements appropriate for RESMOD). We recognize that collection systems allow free text observations, however the IG states that STRESC contains only the base observation without modifiers.
- Avoid embedding UTF8 characters in SEND submissions (e.g., some compounds have Greek letters embedded in them, diacritical marks), we should limit the contributions to 7-bit ASCII
- It is important to remember that the variables, --ANTREG, --LAT, --DIR, and --PORTOT describe the specimen examined. The IG recognizes a difference between the specimen examined and anatomical regions included in the result (ORRES). For example, EYE may be the specimen examined while the result may describe an abnormality in the retina.

Learning Points: MI, MA Domains (cont.)

- Do not provide information in RESMOD or in other SUPPMI or SUPPTF variables that is already contained in another result variable, such as --SEV, --STRESC, --RESCAT, or --DTHREL other than ORRES. This was shown in examples in SENDIGV3.0 (and has been clarified in assumptions of SENDIGV3.1, to state that: MIRESMOD must be populated if one or more modifiers were part of the result in MIORRES and not otherwise reported in the modifier variables part of the MI domain structure, e.g., MISEV).

- The preference for the FDA is not to include "present" in the supplemental record, SUPPMA or SUPPMI, since the presence of a record in the parent domain, MA or MI, already indicates that the finding is present (add enhanced wording V3.1)

Learning Point: TX Domain

- Recovery and/or toxicokinetic animals should typically be presented in separate Trial Sets from the main arm. Trial Sets should be defined to contain animals of both sexes if all other experimental parameters are the same.*

* Excerpt from Study Data Technical Conformance Guide v3.3 dated March 2017, available on the FDA website, "Study Data Standards Resources", https://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm#guides

Learning Points: nSDRG, Define file

- Make use of the nSDRG - the reviewers are using it as we expected! (e.g., many examples of the reviewer looking to the nSDRG first when they just weren't sure or had a question.)
- Make sure your define files are created correctly (e.g., a few examples were seen in the pilot where variable labels and variable types weren't matching the IG)
- The Validator needs to be specifically run on the define file.
- Define v1.0 may cause validator messages that will need to be explained in the nSDRG. We recommend using the Define v2.0 (with errata) or later. (Define v1.0 is sunsetting in the Data Standards Catalog)
- There is a mismatch between the Define examples in the SENDIG and Define standards CT code list (e.g. Special-Purpose) for class attributes. This mismatch is in both SENDIG v3.0 and SENDIG v3.1 and has been addressed via errata (see SENDIG public wiki site for errata*). The define standard is expecting the class from the CT code list for General Observation Class.

* SENDIG V3.1 Errata public wiki site: https://wiki.cdisc.org/display/PUB/SENDIG+v3.1+Errata At the time of this writing, errata for SENDIG V3.0 had not yet been published to the CDISC public wiki.

Learning Points: nSDRG, Define file (cont.)

- A learning point for the FDA was to adjust their tool, Janus Nonclinical, to ignore date and datetime variable lengths in the define file.
- All extensions to codelists should include a meaningful value in the "description" field in the define.xml file and the "meaning" column in the nSDRG.
- File naming in conformance with CDISC and eCTD structure is critical for a submission to pass the technical rejection criteria. File naming as per the eCTD structure requires filenames to be in lower case (Note: The Study Data Reviewers Guide should be "nsdrg.pdf" when placed in the ICH eCTD IG v4.0* structure.)
- Ensure links to stylesheets and other files in the define.xml are correct and functional. A stylesheet should be provided in the same directory location as the study data.
- Supply valid length attributes for data types when required by the define standard.
- Each variable should have only one ItemDef element within each domain for which the variable exists.

* Section 5.2 of ICH eCTD Implementation Guide http://estri.ich.org/new-eCTD/index.htm

Learning Points: multiple domains

- It is known that there can be differences between the dataset visualization and reporting values based on the precision and rounding for SEND data versus reporting. This is a situation that preparers and receivers need to be aware of. These differences should be noted in the nSDRG and where possible, especially in PP (e.g., AUC) that the method and timing of rounding should be defined, if available, for re-creation purposes.

- In some of the pilot submissions, ECG and PP data were collected but not submitted in SEND. The FDA would expect to see these in SEND datasets for studies within the scope of the SEND standard. Further clarification on telemetry is needed.

Learning Points: multiple domains (cont.)

- Expected variables need to be in the dataset even if NULL (related to this: If an expected variable is entirely NULL, this needs to be explained in the define file)
- If a value in the data is not in the codelist, the extension to the codelist needs to be in the define file and should also be submitted to CDISC/CT via the NCI new term form: https://ncitermform.nci.nih.gov/ncitermform/?version=cdisc
- Be very clear on any terminology used - clear mapping from source system to reporting to SEND dataset. It is very useful to the receiver to describe this in the nSDRG (especially MISEV mappings when the collected terms are different in any way from the standard lists).
- Ensure that Variable labels match what is in the IG
- Variables within a domain should be submitted in the order as stated in the IG

Participant Summary Statements

- Piloting is of great value but also comes at a great cost (effort)
- Any future pilots need to be smaller and focused with clear criteria for participation
- Establishing the correct number of participants is important to represent the diversity in the industry (e.g. diverse implementations) while also controlling the scope of work and effort. Define the data that is needed carefully
- It is important to have a liaison identified early who will aggregate all specific feedback into an anonymized set for review by the group