Data transfer specifications template

Author: Jonathan Jaeger

First release: 20-Jan-2023 - Last modification: 09-Mar-2023

This document is a template for data transfer specifications. It contains the backbone structures and instructions to fill out the sections. Instructions are in italic blue and hidden when document is printed out. Remove this front page before finalizing the document.

Company logo

Data transfer specifications

|  |  |
| --- | --- |
| **Protocol title:** |  |
| **Protocol number:** |  |
| **Data supplier name:** |  |
| **DTS Version (date):** |  |
| **Sponsor:** | Company name  Address |
| **Investigational therapy:** |  |
| **Development phase:** |  |
| **EudraCT number:** |  |
| **IND number:** |  |

Property of Company Name   
Confidential  
May not be used, divulged, published, or otherwise disclosed without the consent of Company Name

Signature page

Instructions: Update signature page as relevant.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author** | | | | |
|  |  |  |  |  |
| **Name, Degree(s)**  **Title** |  | **Date** |  | **Signature** |
| **Reviewer** | | | | |
|  |  |  |  |  |
| **Name, Degree(s)**  **Title** |  | **Date** |  | **Signature** |
| **Approver** | | | | |
|  |  |  |  |  |
| **Name, Degree(s)**  **Title** |  | **Date** |  | **Signature** |

Data transfer specification amendment history

Instructions: Complete the data transfer specifications amendment history as relevant.

Document history

|  |  |
| --- | --- |
| **Data transfer specifications version** | **Date** |
| Amendment N | DD-Mmm-YYYY |
| Amendment N-1 | DD-Mmm-YYYY |
| … | DD-Mmm-YYYY |
| Amendment 1 | DD-Mmm-YYYY |
| Original | DD-Mmm-YYYY |

Amendment N (DD-Mmm-YYYY)

Amendment rationale

Changes to the previous version of the data transfer specifications

Amendment N-1 (DD-Mmm-YYYY)

Amendment rationale

Changes to the previous version of the data transfer specifications

Amendment 1 (DD-Mmm-YYYY)

Amendment rationale

Changes to the previous version of the data transfer specifications

Table of contents

[Signature page 3](#_Toc125122594)

[Data transfer specification amendment history 4](#_Toc125122595)

[Document history 4](#_Toc125122596)

[Amendment N (DD-Mmm-YYYY) 4](#_Toc125122597)

[Amendment N-1 (DD-Mmm-YYYY) 4](#_Toc125122598)

[Amendment 1 (DD-Mmm-YYYY) 4](#_Toc125122599)

[List of abbreviations 6](#_Toc125122600)

[1. Contact information 7](#_Toc125122601)

[1.1. 7](#_Toc125122602)

[1.2. 7](#_Toc125122603)

[2. Data transfer specifications 8](#_Toc125122604)

[3. Study timelines 9](#_Toc125122605)

[4. Delivery requirements 10](#_Toc125122606)

[5. Transfer file specifications 11](#_Toc125122607)

[5.1. Data file structure 11](#_Toc125122608)

[5.2. Test definition 17](#_Toc125122609)

[5.3. Visit definition 18](#_Toc125122610)

List of abbreviations

Instructions: Provide list of all abbreviations used in the document. Abbreviations and acronyms are defined where first used in the document.

| **Acronym** | **Definition** |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |

# Contact information

Instructions: Complete information with regards to the contact information. Change name of the sponsor and the electronic data supplier name as relevant. If needed, change the layout to landscape.

## *<Sponsor>*

|  |  |  |  |
| --- | --- | --- | --- |
| Role | Data management | Programming | <if applicable, add as many as necessary> |
| Contact name |  |  |  |
| Title |  |  |  |
| Address |  |  |  |
| Phone |  |  |  |
| Email |  |  |  |

## *<Electronic data supplier name>*

|  |  |  |  |
| --- | --- | --- | --- |
| Role | <Role 1> | <Role 2> | <if applicable, add as many as necessary> |
| Contact name |  |  |  |
| Title |  |  |  |
| Address |  |  |  |
| Phone |  |  |  |
| Email |  |  |  |

# Data transfer specifications

Instructions: This section covers information related to the data type, naming conventions, transfer schedule, transfer format, transfer type and transfer medium.

|  |  |
| --- | --- |
| Data type: | Laboratory - *<specify the type of laboratory data>*  MRI  ECG  Imaging  Protocol deviation  Other - *<specify the type of data>* |
| File Naming | <PROTOCOL>\_<Type of data>\_<External Data Supplier>\_Data\_Transfer\_YYYYMMDD (where YYYYMMDD is the date of the file transfer in ISO8601 format) |
| Transfer schedule: | Annually - *<specify expected date, e.g., every first Thursday of the year>*  Bi-annually- *<specify expected date, e.g., every first Wednesday of the semester>*  Quarterly- *<specify expected date, e.g., every first Tuesday of the quarter>*  Monthly- *<specify expected date, e.g., every first Monday of the month>*  Weekly - *<specify expected date, e.g., every Monday>*  Upon request |
| Transfer format: | Apache Parquet  SAS datasets (V8 .sas7bdat files)  SAS transport file in .xpt format (CPORT)  Oracle export file (dmp file)  .csv  .xlsx  JSON  Apache Parquet  Other - *<specify the format>* |
| Transfer type: | Cumulative  Incremental |
| Transfer medium | *<Sponsor>* SFTP server  *<Electronic data supplier>* SFTP server  Email with password protection – Data and password have to be shared into two different emails. Password is to be changed for each and every transfer  Other - *<specify the transfer medium>* |

# Study timelines

Instructions: Complete the text as appropriate.

Interim data transfers will occur *<provide high level description of the transfer period>*.

Final data transfer will occur *<provide description of when the last transfer will be done>*.

# Delivery requirements

Instructions: General text not to be modified, except information marked as such.

* Data Transfer will include *< describe high level the content of the data package>*.
* A cover letter must be submitted with each data transfer stating the following information:
  + Sponsor
  + Protocol number
  + Data supplier name
  + Data transfer specification version number
  + Name of files transmitted
  + File transfer generation date
  + Transfer date
  + Associated number of records in each file transmitted
* All transfers must be in the format of the approved test transfer/DTS
* The sponsor may reject any data transfer that does not meet agreed specifications.

# Transfer file specifications

## Data file structure *<example for laboratory data>*

Instructions: Adapt the below table as relevant. It is expected that variables names, label, type, codelist are aligned with CDISC requirements. In case of blinded study, indicate if information is to be kept blinded.

| **Variable Name** | **Variable Label** | **Type** | **Controlled Terms, Codelist or Format** | **CDISC Notes** | **Blinded (Y/N)** |
| --- | --- | --- | --- | --- | --- |
| STUDYID | Study Identifier | Char | <protocol number> | Unique identifier for a study. |  |
| DOMAIN | Domain Abbreviation | Char | LB | Two-character abbreviation for the domain.  If not laboratory data, please adapt the domain and the variable names |  |
| USUBJID | Unique Subject Identifier | Char | 100101 | Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product. |  |
| LBSEQ | Sequence Number | Num |  | Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number. |  |
| LBGRPID | Group ID | Char |  | Used to tie together a block of related records in a single domain for a subject. |  |
| LBREFID | Specimen ID | Char |  | Internal or external specimen identifier. Example: Specimen ID. |  |
| LBSPID | Sponsor-Defined Identifier | Char |  | Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor’s operational database. Example: Line number on the Lab page. |  |
| LBTESTCD | Lab Test or Examination Short Name | Char | As per LBTESTCD standard CDISC terminology (see Lab Test section below) | Short name of the measurement, test, or examination described in LBTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in LBTESTCD cannot be longer than 8 characters, nor can it start with a number (e.g.”1TEST”). LBTESTCD cannot contain characters other than letters, numbers, or underscores. Examples: ALT, LDH. |  |
| LBTEST | Lab Test or Examination Name | Char | As per LBTEST standard CDISC terminology (see Lab Test section below) | Verbatim name of the test or examination used to obtain the measurement or finding. Note any test normally performed by a clinical laboratory is considered a lab test. The value in LBTEST cannot be longer than 40 characters. Examples: Alanine Aminotransferase, Lactate Dehydrogenase. |  |
| LBCAT | Category for Lab Test | Char | (see Lab Test section below) | Used to define a category of related records across subjects. Examples: such as HEMATOLOGY, URINALYSIS, and CHEMISTRY. |  |
| LBSCAT | Subcategory for Lab Test | Char | \* | A further categorization of a test category such as DIFFERENTIAL,  COAGULATON, LIVER FUNCTION, ELECTROLYTES. |  |
| LBORRES | Result or Finding in Original Units | Char |  | Result of the measurement or finding as originally received or collected. |  |
| LBORRESU | Original Units | Char | As per UNIT standard CDISC terminology | Original units in which the data were collected. The unit for LBORRES. Example: g/L. |  |
| LBORNRLO | Reference Range Lower Limit in Orig Unit | Char |  | Lower end of reference range for continuous measurements in original units.  Should be populated only for continuous results. |  |
| LBORNRHI | Reference Range Upper Limit in Orig Unit | Char |  | Upper end of reference range for continuous measurements in original units.  Should be populated only for continuous results. |  |
| LBSTRESC | Character Result/Finding in Std Format | Char | As per LBSTRESC standard CDISC terminology | Contains the result value for all findings, copied or derived from LBORRES in a standard format or standard units. LBSTRESC should store all results or findings in character format; if results are numeric, they should also be stored in numeric format in LBSTRESN. For example, if a test has results “NONE”, “NEG”, and “NEGATIVE” in LBORRES and these results effectively have the same meaning, they could be represented in standard format in LBSTRESC as “NEGATIVE”. |  |
| LBSTRESN | Numeric Result/Finding in Standard Units | Num |  | Used for continuous or numeric results or findings in standard format; copied in numeric format from LBSTRESC. LBSTRESN should store all numeric test results or findings. |  |
| LBSTRESU | Standard Units | Char | As per UNIT standard CDISC terminology | Standardized unit used for LBSTRESC or LBSTRESN. |  |
| LBSTNRLO | Reference Range Lower Limit-Std Units | Num |  | Lower end of reference range for continuous measurements for LBSTRESC/LBSTRESN in standardized units. Should be populated only for continuous results. |  |
| LBSTNRHI | Reference Range Upper Limit-Std Units | Num |  | Upper end of reference range for continuous measurements in standardized units. Should be populated only for continuous results. |  |
| LBSTNRC | Reference Range for Char Rslt-Std Units | Char |  | For normal range values that are character in ordinal scale or if categorical ranges were supplied (e.g., “-1 to +1”, “NEGATIVE TO TRACE”). |  |
| LBNRIND | Reference Range Indicator | Char | As per NRIND standard CDISC terminology | 1. Indicates where the value falls with respect to reference range defined by LBORNRLO and LBORNRHI, LBSTNRLO and LBSTNRHI, or by LBSTNRC. Examples: NORMAL, ABNORMAL, HIGH, LOW. 2. Sponsors should specify in the study metadata (Comments column in the define.xml) whether LBNRIND refers to the original or standard reference ranges and results. 3. Should not be used to indicate clinical significance. |  |
| LBSTAT | Completion Status | Char | As per ND standard CDISC terminology | Used to indicate exam not done. Should be null if a result exists in LBORRES. |  |
| LBREASND | Reason Test Not Done | Char |  | Describes why a measurement or test was not performed such as BROKEN EQUIPMENT, SUBJECT REFUSED, or SPECIMEN LOST. Used in conjunction with LBSTAT when value is NOT DONE. |  |
| LBNAM | Vendor Name | Char |  | The name or identifier of the laboratory that performed the test. |  |
| LBLOINC | LOINC Code | Char | \* | 1. Dictionary-derived LOINC Code for LBTEST. 2. The sponsor is expected to provide the dictionary name and version used to map the terms utilizing the define.xml external codelist attributes |  |
| LBSPEC | Specimen Type | Char | As per SPECTYPE standard CDISC terminology | Defines the type of specimen used for a measurement. Examples: SERUM, PLASMA, URINE. |  |
| LBSPCCND | Specimen Condition | Char | As per SPECCOND standard CDISC terminology | Free or standardized text describing the condition of the specimen e.g. HEMOLYZED, ICTERIC, LIPEMIC etc. |  |
| LBMETHOD | Method of Test or Examination | Char | As per METHOD standard CDISC terminology | Method of the test or examination. Examples: EIA (Enzyme Immunoassay), ELECTROPHORESIS, DIPSTICK |  |
| LBBLFL | Baseline Flag | Char | As per NY standard CDISC terminology | Indicator used to identify a baseline value. The value should be “Y” or null. |  |
| LBFAST | Fasting Status | Char | As per NY standard CDISC terminology | Indicator used to identify fasting status such as Y, N, U, or null if not relevant. |  |
| LBDRVFL | Derived Flag | Char | As per NY standard CDISC terminology | Used to indicate a derived record. The value should be Y or null. Records that represent the average of other records, or do not come from the CRF, or are not as originally received or collected are examples of records that might be derived for the submission datasets. If LBDRVFL=Y, then LBORRES may be null, with LBSTRESC, and (if numeric) LBSTRESN having the derived value. |  |
| LBTOX | Toxicity | Char | \* | Description of toxicity quantified by LBTOXGR. The sponsor is expected to provide the name of the scale and version used to map the terms, utilizing the define.xml external codelist attributes. |  |
| LBTOXGR | Standard Toxicity Grade | Char | \* | Records toxicity grade value using a standard toxicity scale (such as the NCI CTCAE). If value is from a numeric scale, represent only the number (e.g., “2” and not “Grade 2”). The sponsor is expected to provide the name of the scale and version used to map the terms, utilizing the define.xml external codelist attributes. |  |
| VISITNUM | Visit Number | Num | See Visit definition below | 1. Clinical encounter number. 2. Numeric version of VISIT, used for sorting. |  |
| VISIT | Visit Name | Char | See Visit definition below | 1. Protocol-defined description of clinical encounter 2. May be used in addition to VISITNUM and/or VISITDY |  |
| LBDTC | Date/Time of Specimen Collection | Char | ISO 8601 |  |  |
| LBENDTC | End Date/Time of Specimen Collection | Char | ISO 8601 |  |  |

<add as necessary>

## Test definition *<example for laboratory>*

Instructions: Adapt the below table as relevant.

| Protocol Lab Test Name | Electronic Data Supplier Lab Test Name | LBTESTCD Lab Test or Examination Short Name | Lab Test Code (LBTEST standard CDISC terminology) | Category for Lab Test (LBCAT) | Subcategory for Lab Test (LBSCAT) | Specimen type (LBSPEC) | Method of Test or Examination (LBMETHOD) |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | WBC | WBC | Leukocytes | HEMATOLOGY |  | BLOOD | FLOW CYTOMETRY |
|  |  |  |  |  |  |  |  |

<add as necessary>

## Visit definition

Instructions: Adapt the below table as relevant.

| Electronic Data Supplier Visit Code | Sponsor Visit Code | Electronic Data Supplier Visit Name | Sponsor Visit Name |
| --- | --- | --- | --- |
|  | 10 |  | Screening |
|  | 20 |  | Baseline |

<add as necessary>

If repeat of a scheduled visit or of an unscheduled visit, a repeated analysis will be labeled as

* Visit Name: previous scheduled visit associated with Unscheduled (e.g. Baseline Unscheduled)
* Visit Number: Previous Visit Number associated with .1, .2, etc…in a chronological order (e.g. 20.1)