**CONSIDER Statement:**

**Consolidated Recommendations for Sharing Individual Participant Data from Human Clinical Studies**

This document provides a checklist and a set of recommendations to guide principal investigators (PIs), study team members and data sharing platform representatives in optimal sharing of individual participant data (IPD) from human clinical studies. We use the term study to include interventional trials and observational studies.

Recommendations are structured by several domains. Checklist below represents a brief summary of recommendations. The details link leads to elaborated descriptions of each item.

We welcome feedback to any checklist item at craig.mayer2@nih.gov

The acronym is loosely derived from letters contained in the title: **CONS**olidated **RE**commendations for sharing **I**ndividual participant **D**ata). Letters E and R are re-ordered to create a more memorable acronym.

# Data Format

## Share person table in CDISC or OMOP format

[positive-example](#_Positive_Example), challenge-example, [details](#_Details)

## For improved integration of research and routine healthcare data (e.g., from Electronic Health Record system or healthcare billing data), group relevant data elements into relevant data domains (e.g., medication history, laboratory results history, medical procedure history)

positive example

challenge example

details  
Emergence of several common data models (CDMs) for healthcare data shows that there is a common way to organize clinical data. For example, rather than considering numerical value as test result and unit in which the value is expressed as two separate data elements, a review of several CDMs shows that these are typically grouped into a single data row.

# Data Sharing

## Use ClinicalTrials.gov registry to document your study

[positive-example](#_Positive_Example_1)

## Do not limit study metadata to the legally required elements. Also populate optional elements (such as data sharing metadata)

challenge-example

## On ClinicalTrials.gov, if you answer Yes to share\_ipd\_data, do not leave the data\_sharing\_plan text empty

[challenge-example](#_Missing_sharing_plan): NCT02756208 and NCT03275701

# Data dictionary

## Provide data dictionary documentation separate from de-identified individual participant data. Since it contains no participant level data, do not require local ethical approval as a condition of releasing the data dictionary (avoid a request-wall).

positive-example, details  
details  
If DD does contain important intellectual property (IP), consider creating a smaller list of DEs that do not contain any IP and release this limited subset of DEs without employing a request-wall.

## Provide data dictionary in a single, machine-readable file.

[challenge-example](#_Scattered_data_dictionary), details

NCT01233531 has 17 data dictionary files, which includes documents in different formats and document types. Also includes identical file names that represent dictionaries for different data based on visit type and study population group.

details  
This simplifies machine processing of available study data. Using a single file approach also ensures that each file (if scattered across multiple) uses the same structure (e.g., DE label, DE data type, DE permissible values [for categorical DEs])

## For each data element, provide a data type (such as numeric, date, string)

details  
Specifying data type helps computers to process the information properly. Data type also helps with semantic matching to corresponding CDEs. For example, date of death data type is stated as date (not as character).

challenging example

NCT00046280 does not provide data type

## For categorical data elements, provide a list of permissible values

details  
For example, for educational level data element, it is important to know what possible values were considered during data collection. While it is possible to discover those permissible values from IPD, if some values were never applicable to any of the subjects, the reverse-engineered permissible value list will be incomplete. In terms of standards, CDISC ODM and REDCap provide a mechanism to list permissible values.

challenging example

NCT00000590 does not provide permissible vales for categorical variables in the data dictionary

positive example

NCT00683579 provides permissible values and definitions for the values associated with categorical data elements.

## Link data elements (including individual permissible values) to applicable routine healthcare terminologies or research common data elements

positive example  
AllOfUs study links DEs to SNOMED CT and LOINC codes. For example, Body Mass Index data element is clearly linked to LOINC code of 39156-5 (for BMI)

challenge example

NCT01751646 has no link to any standardized vocabulary

details  
This recommendation does not mean that all DEs (and permissible values) are linked to a terminology or applicable standard. Only where a relevant code exist (or is easy to find and reference) this link is recommended.

# Data de-identification

## Provide data de-identification notes

positive-example, challenge-example, details

# Data format

## Use consistent relative time (start counting at day 1, not 0)

Details

When using relative time, specify index event with datetime granularity. Refer to the first day as relative day 1. Do not use day 0 as a relative date. For example if the index event is signing of informed consent and it was signed at 10:31am on March 10, 2011, the index date-time is midnight of March 10, 2011. In relative time, of an event on the next day (on March 11) at 11:15am, would have relative time of Day 2, 11:15am. (for analogous discussion in astronomical data see <https://en.wikipedia.org/wiki/Sol_(day_on_Mars)#Usage_in_Mars_landers>)

# Abbreviations

DD = data dictionary

IPV=Individual Participant Data

# Details

## Data format

### Utilize ClinicalTrials.gov fields for uploading study protocol, empty case report forms, statistical analysis plan and link to

Note: Example trial is here (TODOVH)

### Provide basic summary results using results registry component of Clinicaltrials.gov

### Share Case Report Forms in non-PDF, machine readable format. For example, REDCap, OpenClinica and several other Electronic Data Capture (EDC) systems allow export into CDISC Operational Data Model (ODM) format for forms. If no cross-platform standard is supported by your EDC, provide CRFs in the platform-specific format.

### If you considered formally defined research Common Data Elements at study design (more common for studies initiated after 2015), provide a spreadsheet file that lists all CDEs utilized by your study. Include unique CDE identifies (e.g., PhenX VariableID).

### Use data formats that can be natively loaded (without add-ons) into multiple statistical platforms (e.g., prefer comma/tab separated values (.CSV) files to SAS XPT, XLS/XSLX)

### Distinguish string field with free text from string field (with a set of permissible values)

### For pick list questions, provide in data dictionary the options that are possible

Example from NIDA: TODOSG

### Provide a data dictionary in one spreadsheet for all data.

Why: Automated processing of data dictionary

Note: TODOSG NIDA had 65 data files but only 63 data dictionary PDF files. Matching data files with data dictionary requires manual matching.

### Provide a complete data dictionary (all data files must be included)

Note: NIDA trial had 5 data files that were missing a data dictionary file.

### Person table

Share Person table Person table

### Positive Example

Trial <http://clinicaltrials.gov/ct2/show/NCT01612169> IPV data posted on NIDA Data Share (at <https://datashare.nida.nih.gov/study/nidactn0049>) provide file dem.csv provides one row per person with person\_id and basic demographic data. (following CDISC standard).

### Details

OMOP person table specification can be found at <https://github.com/OHDSI/CommonDataModel/wiki/PERSON>

## Positive Example

Trial titled “RC-HIVMAB060-00-AB (VRC01) in People With Chronic HIV Infection Undergoing Analytical Treatment Interruption” is registered at ClinicalTrials.gov under NCT02471326 <https://clinicaltrials.gov/ct2/show/study/NCT02471326>. It allows retrieval of study metadata that is unified across various platform.

## Missing sharing plan description

Out of 69 reviewed HIV trials IPD sharing plans on ClinicalTrials.gov, 2 (NCT02756208 and NCT03275701) left the plan description blank.

## Scattered data dictionary

Trials NCT00005274 and NCT00005274 provided DD in several files. In order to relate those to CDEs, manual processing is required.

## Data Sharing