
CTN-0093 Data De-Identification Notes

And

Database Descriptions

Mar 31, 2023

CTN-0093 Data De-Identification Notes

The data presented here in de-identified form is an analysis dataset, rather than as-collected case report form data. Each study participant is represented by a single row in the dataset.

- Patient ID: All patient IDs were transformed to factless key values bearing no relationship to the collected study data.
- Dates: All dates were converted to numeric values indicating “days on study”, where the date of consent = day 0. A negative value for days on study refers to an event that occurred prior to consent, and a positive value for days on study refers to an event that occurred after consent.
 - Date variables that included a time component will appear as numeric values with fractional days (expressed as decimals), not just integers.
- Initials / Names: Any field designed to record a person’s initials or name was erased.
- Other Identifiers: Sample identifiers and Drug lot numbers were erased.
- Text Fields: Most free text fields long enough to contain dates or narrative were emptied, including
 - Comment fields.
 - Responses derived from “Specify”, “Describe”, “Reason”, or “Other”. For example, other race, other termination reason, reason study drug not given, abnormalities on ECG and lab forms.
 - Narratives regarding interaction with patient, health conditions, treatment progress or follow-up.
 - Medical history and physical exam narrative text.
 - Type of work or job description; narrative descriptions of living situation, illnesses, concerns, drug/alcohol use, feelings/interests.
- Text Field De-identification Exceptions:
 - AE and medical condition terms were *retained* in the database, but detailed descriptions in a separate field providing additional information about adverse events were erased, such as relevant labs, medical history, drug/alcohol use description, probable cause of death.
 - Concomitant medication names and indications were *retained* in the database.
- For a complete listing of the erased text fields, refer to the “CTN0093 Nulled Values” csv file.

See the “CTN0093 Data Dictionary” Excel workbook file for descriptions of the data.

NOTE: A file that discusses the handling of a few data errors is available in the “Study Documents” section of the Datashare page for this study: see “Data Notes for Final Dataset”.