

Collaborative Pediatric Critical Care Research Network

Parent Provider Alliance Study

Overview of Public Use Datasets

This document provides an overview of the general principles used in creating the public use datasets for the Collaborative Pediatric Critical Care Research Network (CPCCRN) Parent Provider Alliance Study (PPA). There are a total of 5 datasets available as both CSV and SAS® (.SAS7BDAT) files (with formats and labels). Accompanying documentation and resources provided include the study protocol, the codebook, the survey booklets, study summary and, for each dataset, a PDF document summarizing variables in the associated dataset (i.e., frequency distributions or descriptive statistics). For SAS software users, an example script to set up the SAS library and apply SAS formats is also provided. The annotated CRF should be referenced frequently during analysis as this is the most complete reference of all variables included in each dataset.

GENERAL PRINCIPLES FOR CREATION OF DATASET

1. The population for the public use dataset is all eligible subjects. See the protocol for a detailed list of inclusion/exclusion criteria and for additional details of study conduct.
2. The datasets are primarily based on raw data as collected by the clinical sites. Categorical variables with small cell counts were collapsed into fewer categories such as parent race, diagnosis, and relation to child. All variables are described in the codebook.
3. Open text fields and other variables have been reviewed for sensitive or identifying information and modified or removed as needed. Dates were recoded to be the number of days since the screened date (date eligibility was determined, not the death date).
4. Within the survey packets developed for the parents to complete, the numbering for response options were changed for the PHQ and PTSD to allow for consistency with the other surveys. Since raw datasets are included in the public use dataset, please note that the PHQ and PTSD item scores will need to be subtracted by one before summing the total scores (ie. The raw PHQ answers for the 9 items are currently '1= Not at all' to '4=Nearly everyday' but need to actually be '0=Not at all' to '3=Nearly everyday' in order to sum correctly between 0 and 27. The raw PTSD answers for the 8 items are '1=Not at all' to '5=Very much' but need to actually range from 0 to 4 in order to sum between 0 and 32)
5. Within each dataset, the PUDID is a masked identifier. The codebook provides information as to the structure of each dataset and the unique identifier for each record.

LIST OF PUBLIC USE DATASETS AVAILABLE

- PPA 6-Month Contact (CONTACT6MONTH_RELATION)
- PPA Survey 6-Month (PPA6MONTHSURVEY2)

- PPA Hospitalization (PPAHOSPITALIZATION)
- PPA 13-Month Contact (PPACONTACT13MONTH_RELATION13)
- PPA Survey 13-Month (PPA13MONTHSURVEY)