

## **Collaborative Pediatric Critical Care Research Network Vent CDS R21 Study Overview of Public Use Datasets**

This document provides an overview of the general principles used in creating the public use datasets for the *Eunice Kennedy Shriver* National Institute for Child Health and Human Development Collaborative Pediatric Critical Care Research Network (CPCCRN) Vent CDS R21 Study. There are a total of 10 datasets available as both CSV and SAS® (.SAS7BDAT) files (with formats and labels). Accompanying documentation and resources provided include the study protocol, the annotated eCRF, 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 eCRF should be referenced frequently during analysis as this is the most complete reference of all variables included in each dataset. Please see the attached Research Data Use Agreement (RDUA) for a description of intended use and disclaimer.

### **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. In addition, two datasets with derived analysis variables have been included. All variables are described in the annotated eCRF.
3. Open text fields and other variables have been reviewed for sensitive or identifying information and modified as needed. Dates were recoded to be the number of days since the subject's first ventilator reading.
4. Within each dataset, the SubjectID is a masked identifier. For datasets with more than one record per subject, SubjectID + ItemGroupRepeatKey is the unique identifier. The annotated eCRF provides information as to the structure of each dataset and the unique identifier for each record.

### **LIST OF PUBLIC USE DATASETS AVAILABLE**

- Ventilator Start Information (ABGVENT)
- Blood Gases Log (ABGVENT\_BLOODGAS)
- SPO2 and End Tidal CO2 Log (ABGVENT\_SPO2)
- Ventilator Settings Log (ABGVENT\_VENT)
- Additional Data Elements (ADDDATA)
- Eligibility (ELIGIBILITY)
- Radiology Information (LABRAD\_CXR)
- Hemoglobin Lab Results (LABRAD\_HEMLABS)
- Subject-level Derived Variables (SUBJECTLEVELDERIVED)
- Ventilator Settings Log Joined with SPO2 and Blood Gases Logs (VENTCHANGES)