Critical Illness Stress-Induced Immune Suppression (CRISIS) Prevention Trial
Overview of Public Use Datasets

This document overviews 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) CRISIS Prevention Trial. There are a total of 29 datasets available as either CSV or SAS files (with formats and labels). Accompanying documentation and resources provided include the study protocol, an overview of the outcome adjudication process, the annotated eCRF, and, for each dataset, a PDF document summarizing variables in that dataset (i.e., frequency distributions or descriptive statistics). For SAS 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.

GENERAL PRINCIPLES FOR CREATION OF DATASET

1. The population for the public use dataset is all randomized patients. Screening information is not included. 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, key analysis variables relating to the primary and secondary outcomes are included. All variables are described in the annotated eCRF.

3. Sensitive or identifying information has been removed throughout the dataset as follows:
    - Dates have been recoded to reflect the number of calendar days from randomization.
    - Limited open text fields are included. When included, these fields were reviewed in detail and any sensitive or identifying information was removed (i.e., entry blanked out).
    - We have suppressed any values with a count < 5 that may have sensitive or identifying information. Notes to this effect are included on the annotated eCRF. For example, for primary diagnosis, the category of *HIV Infection* was combined with *Other*.

4. If variables were collected in multiple locations within the CRF, the public use dataset only includes the final data source utilized for study analyses. In addition, variables collected for study purposes only that would not be relevant to clinical research are not included. Variables not included in the public use dataset do not appear on the annotated eCRF or are identified as *Not included*.

5. Many of the available datasets include only one record per patient (unique identifier *subjectID*). Other datasets are relational, that is, may have more than one record per patient. 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

- Demographics
- Randomization
- Baseline and Pre-Dosing
- Baseline - Chronic Diagnoses
- Physical Examination
- Daily Data / Follow Up
- Daily Data - Additional Therapies
- Pediatric Logistic Organ Dysfunction (PELOD) Score
- Organ Failure Index (OFI)
- Daily Hematology Results
- Daily Serum Chemistry Results
- Miscellaneous Laboratory Values
- Final Status (includes Day 28 mortality information)
- Final Status – Discontinued Study Medications
- Withdrawal of Consent
- Adverse Events Log
- Concomitant Medications Log
- Diagnostic Testing for Infections Log
- Parenteral Study Drug Administration
- Metoclopramide or Placebo Study Drug Administration Log
- Enteral Study Drug Administration
- Zinc or Placebo Study Drug Administration Log
- Selenium or Placebo Study Drug Administration Log
- Glutamine or Placebo Study Drug Administration Log
- Antibiotic Log
- Endpoint Summary
- Endpoint Summary - Nosocomial Infection Events Log
- Endpoint Summary - Nosocomial Sepsis Events Log
- Mayo Laboratory Data