General Descriptors

As noted, these six descriptors are collected at the time the patient is admitted to the ICU. Their associated time-stamps are set to 00:00 (thus they appear at the beginning of each patient's record).

- *RecordID* (a unique integer for each ICU stay)
- *Age* (years)
- *Gender* (0: female, or 1: male)
- *Height* (cm)
- *ICUType* (1: Coronary Care Unit, 2: Cardiac Surgery Recovery Unit,
- 3: Medical ICU, or 4: Surgical ICU)
- *Weight* (kg)*.

The ICUType was added for use in Phase 2; it specifies the type of ICU to which the patient has been admitted.

Time Series

These 37 variables may be observed once, more than once, or not at all in some cases:

- *Albumin* (g/dL)
- *ALP* [Alkaline phosphatase (IU/L)]
- *ALT* [Alanine transaminase (IU/L)]
- *AST* [Aspartate transaminase (IU/L)]
- *Bilirubin* (mg/dL)
- *BUN* [Blood urea nitrogen (mg/dL)]
- *Cholesterol* (mg/dL)
- *Creatinine* [Serum creatinine (mg/dL)]
- *HCT* [Hematocrit (%)]
- *HR* [Heart rate (bpm)]
- *K* [Serum potassium (mEq/L)]
- *Lactate* (mmol/L)
- *Mg* [Serum magnesium (mmol/L)]
- *MAP* [Invasive mean arterial blood pressure (mmHg)]
- *MechVent* [Mechanical ventilation respiration (0:false, or 1:true)]
- *Na* [Serum sodium (mEq/L)]
- *PaCO2* [partial pressure of arterial CO_2 (mmHg)]
- *PaO2* [Partial pressure of arterial O_2 (mmHg)]
- *pH* [Arterial pH (0-14)]
- *Platelets* (cells/nL)
- *RespRate* [Respiration rate (bpm)]
- *SaO2* [O_2 saturation in hemoglobin (%)]
- *SysABP* [Invasive systolic arterial blood pressure (mmHg)]
- *Temp* [Temperature (°C)]

- *DiasABP* [Invasive diastolic arterial blood pressure (mmHg)]
- *FiO2* [Fractional inspired O_2 (0-1)]
- *GCS* [Glasgow Coma Score (3-15)]
- *Glucose* [Serum glucose (mg/dL)]
- *HCO3* [Serum bicarbonate (mmol/L)]
- *NIDiasABP* [Non-invasive diastolic arterial blood pressure (mmHg)]
- *NIMAP* [Non-invasive mean arterial blood pressure (mmHg)]
- *NISysABP* [Non-invasive systolic arterial blood pressure (mmHg)]
- *TropI* [Troponin-I (µg/L)]
- *TropT* [Troponin-T (µg/L)]
- *Urine* [Urine output (mL)]
- *WBC* [White blood cell count (cells/nL)]
- *Weight* (kg)*

The time series measurements are recorded in chronological order within each record, and the associated time stamps indicate the elapsed time since admission to the ICU. Measurements may be recorded at regular intervals ranging from hourly to daily, or at irregular intervals as required. Not all time series are available in all cases.

In a few cases, such as blood pressure, different measurements made using two or more methods or sensors may be recorded with the same or only slightly different time-stamps. Occasional outliers should be expected as well.

* Note that *Weight* is both a general descriptor (recorded on admission) and a time series variable (often measured hourly, for estimating fluid balance).

Outcome-related Descriptors

The outcome-related descriptors are kept in a separate CSV text file for each of the three record sets; as noted, only the file associated with training set A is available to participants. Each line of the outcomes file contains these descriptors:

- *RecordID* (defined as above)
- *SAPS-I score* (Le Gall et al., 1984)
- *SOFA score* (Ferreira et al., 2001)
- *Length of stay* (days)
- *Survival* (days)
- *In-hospital death* (0: survivor, or 1: died in-hospital)

The *Length of stay* is the number of days between the patient's admission to the ICU and the end of hospitalization (including any time spent in the hospital after

discharge from the ICU). If the patient's death was recorded (in or out of hospital), then *Survival* is the number of days between ICU admission and death; otherwise, *Survival* is assigned the value -1. Since patients who spent less than 48 hours in the ICU have been excluded, *Length of stay* and *Survival* never have the values 0 or 1 in the challenge data sets. Given these definitions and constraints,

Survival > *Length of stay* ⇒ Survivor
Survival = -1 ⇒ Survivor
2 ≤ *Survival* ≤ *Length of stay* ⇒ In-hospital death