BOLD, a blood-gas and oximetry linked dataset

Data Dictionary

Demographics

subject_id: Describes a unique subject. This is unique per component dataset and mapped directly to the equivalent term in each. A subject may have multiple admissions, denoted by hospital_admission_id. Same as in the original database.

hospital_admission_id: Describes a unique hospital admission. This is unique per component dataset and mapped directly to the equivalent term in each. Same as in the original database.

icustay_id: Describes a unique ICU admission. This is unique per component dataset and mapped directly to the equivalent term in each. Each hospital admission may have multiple ICU stays. Same as in the original database.

unique_subject_id: Describes a unique subject. Guarantees that no subjects coming from different databases have the same identifier.

unique_hospital_admission_id: Describes a unique hospital admission. Guarantees that no subjects coming from different databases have the same identifier.

unique_icustay_id: Describes a unique ICU admission. Guarantees that no subjects coming from different databases have the same identifier.

source_db: Labeled as *mimic_iii*, *mimic_iv*, *eicu* to distinguish between component datasets.

race_ethnicity: Harmonized race and ethnicity .

Hospital characteristics

hospitalid: Unique hospital ID. Beth Israel Deaconess (MIMIC-III, -IV) was denoted as 9999.

numbedscategory: Hospital size, in numbers of beds. This was recorded as "*< 100*"; "*100-250*"; "*250-500*"; "*≥ 500*" beds for eICU-CRD. For MIMIC-III and -IV (BIDMC), we fixed the value at "*≥ 500" beds*.

teachingstatus: This field marks whether a hospital was identified as a teaching hospital, where a value of 1 implies that it was a teaching hospital. BIDMC has teaching status = 1.

region: Maps to the US census region distribution per eICU. It is either *Midwest, Northeast, South*, or *West*. MIMIC (BIDMC) was set as *Northeast*.

admission_age: Harmonized age on admission. eICU admission age mapped to admission_age. MIMIC-III admission age mapped to admission age. MIMIC-IV admission_age mapped to admission_age.

sex_female: Assigned a value of 1 if the patient is of female sex.

weight_admission: Weight on admission (in kilograms).

height_admission: Height on admission (in centimeters).

BMI_admission: Calculated BMI on admission, based on weight_admission and height_admission. If either weight_admission or height_admission is missing, BMI_admission is missing.

Admission characteristics

datetime_hospital_admit, datetime_hospital_discharge, datetime_icu_admit, datetime_icu_discharge: Date and time of hospital and ICU admission (_admit) and discharge (_discharge).

los_hospital, los_ICU: Length of stay for hospital (_hospital) and ICU (_ICU) in days.

in_hospital_mortality: This variable is true if the patient died during the hospital admission, regardless if the patient died during the ICU admission or not.

comorbidity_score_name, comorbidity_score_value: Comorbidity score (either Elixhauser or Charlson), along with score.

ABG data

SaO2_timestamp: Date and time of ABG test.

pH: pH value.

pCO2: Partial pressure of CO2 (mmHg).

pO2: Partial pressure of O2 (mmHg).

SaO2: Arterial oxygen saturation (%).

Carboxyhemoglobin: Percentage of hemoglobin bound to CO (%).

Methemoglobin: Percentage of methemoglobin (%).

Vitals data

SpO2: Pulse oximetry saturation (%).

SpO2_timestamp: Date and time of SpO2 recording.

vitals_heart_rate: Heart rate.

vitals_resp_rate: Respiratory rate.

vitals_mbp_ni, vitals_sbp_ni, vitals_dbp_ni: Mean arterial pressure (MAP), Systolic pressure (SBP), Diastolic pressure (SBP) calculated from noninvasive (cuff) blood pressure.

vitals_mbp_i, vitals_sbp_i, vitals_dbp_i: Mean arterial pressure (MAP), Systolic pressure (SBP), and Diastolic pressure (DBP) calculated from invasive (arterial line) blood pressure.

vitals_tempc: Temperature, from any body measuring site, in Celsius (ºC).

Labs

Complete Blood Count (CBC)
cbc_wbc: White blood cell. (10^9/L)

cbc_hemoglobin: Measured hemoglobin (g/L).

cbc_hematocrit: Measured hematocrit. (%)

cbc_platelet: Measured platelet count.(10^9/L)

cbc_mch: Measured mean corpuscular hemoglobin. (pg)

cbc_mchc: Measured mean corpuscular hemoglobin concentration. (g/L)

cbc_mcv: Measured mean corpuscular volume.(fL)

cbc_rbc: Measured red blood cells (RBC). (10^{12}/L)

cbc_rdw: Measured RBC distribution width. (%)

Coagulation labs
coag_fibrinogen: Measured fibrinogen.

coag_pt: Measured prothrombin time. (s)

coag_inr: Measured international normalized ratio.

coag_ptt: Measured partial thromboplastin time. (s)

Basic Metabolic Panel (BMP)

bmp_sodium: Measured sodium levels. (mmol/L)

bmp_potassium: Measured potassium levels. (mmol/L)

bmp_chloride: Measured chloride levels. (mmol/L)

bmp_bicarbonate: Measured bicarbonate levels. (mg/dL)

bmp_bun: Measured blood urea nitrogen levels. (mg/dL)

bmp_creatinine: Measured creatinine levels. (mg/dL)

bmp_glucose: Measured glucose levels. (mg/dL)

bmp_aniongap: Measured anion gap. (mmol/L)

bmp_calcium: Measured calcium levels. (mg/dL)

bmp_lactate: Measured lactate levels. (mmol/L)

Hepatic function panel (HFP)

hfp_alt: Measured alanine aminotransferase (ALT) levels. (U/L)

hfp_alp: Measured alkaline phosphatase (ALP) levels. (U/L)

hfp_ast: Measured aspartate aminotransferase (AST) levels. (U/L)

hfp_bilirubin_total: Measured total bilirubin levels. (mg/dL)

hfp_bilirubin_direct: Measured direct bilirubin levels. (mg/dL)

hfp_albumin: Measured albumin levels. (g/dL)

Other labs (enzyme)

others_ck_cpk: Measured creatine kinase (CK) levels, also known as creatine phosphokinase (CPK). (U/L)

others_ck_mb: Measured creatine kinase MB (CK-MB) levels. (U/L)

others_ld_ldh: Measured lactate dehydrogenase (LDH) levels. (U/L)

SOFA scores

sofa_past_overall_24hr: SOFA score, calculated from component values below, measured in the hour window prior to the ABG.

sofa_past_coagulation_24hr, sofa_past_liver_24hr, sofa_past_cardiovascular_24hr, sofa_past_cns_24hr, sofa_past_renal_24hr: SOFA score components, with highest value for each component in the past 24 hours. The hour window just prior to the hour window containing the ABG is recorded here to characterize baseline patient status.

sofa_future_overall_24hr: SOFA score, calculated from component values below, measured 24 hours after the ABG window.

sofa_future_coagulation_24hr, sofa_future_liver_24hr, sofa_future_cardiovascular_24hr, sofa_future_cns_24hr, sofa_future_renal_24hr: SOFA score components, with highest value for each component in the 24 hours after the ABG to characterize the 24 hour impact of discrepancies.