2) Data Mapping

Dataset Overview

This study uses the EPIC data tables from the Medical Informatics Operating Room Vitals and Events Repository (MOVER), which contains electronic medical record data on surgeries performed at the University of California, Irvine (UCI) Medical Center from 2017 to 2022, comprising records for 39,685 patients and 64,354 surgeries.

Identifiers and unit of analysis

The dataset includes two identifiers: MRN (a patient-level identifier) and LOG_ID (an encounter-level identifier). A single patient (MRN) may have multiple surgical encounters, resulting in multiple LOG_ID values. Because postoperative hypoxemia and perioperative variables are encounter-specific, the analytic unit for this study is the surgical encounter, represented by LOG_ID.

Data scope and construct mapping

We will use four EPIC tables—Patient Information, Patient Labs, Patient Procedure Events, and Patient Postoperative Complications—to address our research questions on predictors of postoperative hypoxemia. These tables collectively provide encounter-level demographics, preoperative test results, perioperative timestamps for the surgical encounter, and the postoperative complications used to define the hypoxemia outcome. For preoperative laboratory predictors, we focus on hematocrit and leukocyte (white blood cell) measurements.

Variable mapping

Table 1 summarizes how each construct is represented in the available data tables, including variable names and measurement types.

Construct	Table	Columns	Level	Type
Postoperative hypoxemia (outcome)	Patient Postoperative Complications	SMRTDTA_ELEM_VALUE	Encounter	Binary
Demographics (age, sex, BMI)	Patient Information	BIRTH_DATE, SEX, HEIGHT, WEIGHT	Encounter	Numeric / Categorical
Lab test result (Hematocrit, Leukocytes)	Patient Labs	Lab.Code, Observation.Value, Measurement.Units, Collection.Datetime	Encounter-lab record	Numeric
Anesthesia timing	Patient Procedure Events	EVENT_DISPLAY_NAME, EVENT_TIME	Encounter-event record	Datetime

Table 1. Construct-to-variable mapping (encounter key: LOG_ID).

Note: Collection.Datetime (Patient Labs) and EVENT_TIME (Patient Procedure Events) are alignment timestamps used to relate laboratory measurements to the surgical timeline (e.g., identifying the most recent test results before anesthesia start). EVENT_DISPLAY_NAME is used to identify the anesthesia start event. These fields will not be included as predictors in the final analytic table.

Together, these mapped fields support constructing an encounter-level dataset indexed by LOG_ID, containing the hypoxemia outcome, demographics (age/sex/BMI) and preoperative hematocrit and leukocyte measurements.