**Overview of the MIMIC-III data**

MIMIC is a relational database containing tables of data relating to patients who stayed within the intensive care units at Beth Israel Deaconess Medical Center. A table is a data storage structure which is similar to a spreadsheet: each column contains consistent information (e.g., patient identifiers), and each row contains an instantiation of that information (e.g. a row could contain the integer 340 in the patient identifier column which would imply that the row’s patient identifier is 340).

The tables are linked by identifiers which usually have the suffix “ID”. For example HADM\_ID refers to a unique hospital admission and SUBJECT\_ID refers to a unique patient. One exception is ROW\_ID, which is simply a row identifier unique to that table.

Tables pre-fixed with “D\_” are dictionaries and provide definitions for identifiers. For example, every row of OUTPUTEVENTS is associated with a single ITEMID which represents the concept measured, but it does *not* contain the actual name of the drug. By joining OUTPUTEVENTS and D\_ITEMS on ITEMID, it is possible to identify what concept a given ITEMID represents.

List of tables

The following tables are used to define and track patient stays:

* **ADMISSIONS**: Every unique hospitalization for each patient in the database (defines HADM\_ID)
* **CALLOUT**: Information regarding when a patient was cleared for ICU discharge and when the patient was actually discharged
* **ICUSTAYS**: Every unique ICU stay in the database (defines ICUSTAY\_ID)
* **PATIENTS**: Every unique patient in the database (defines SUBJECT\_ID)
* **SERVICES**: The clinical service under which a patient is registered
* **TRANSFERS**: Patient movement from bed to bed within the hospital, including ICU admission and discharge

Each ICUSTAY\_ID corresponds to a single HADM\_ID and a single SUBJECT\_ID. Each HADM\_ID corresponds to a single SUBJECT\_ID. A single SUBJECT\_ID can correspond to multiple HADM\_ID (multiple hospitalizations of the same patient), and multiple ICUSTAY\_ID (multiple ICU stays either within the same hospitalization, or across multiple hospitalizations, or both).

The following tables contain data collected in the critical care unit:

* **CAREGIVERS**: Every caregiver who has recorded data in the database (defines CGID)
* **CHARTEVENTS**: All charted observations for patients
* **DATETIMEEVENTS**: All recorded observations which are dates, for example time of dialysis or insertion of lines.
* **INPUTEVENTS\_CV**: Intake for patients monitored using the Philips CareVue system while in the ICU
* **INPUTEVENTS\_MV**: Intake for patients monitored using the iMDSoft Metavision system while in the ICU
* **NOTEEVENTS**: Deidentified notes, including nursing and physician notes, ECG reports, imaging reports, and discharge summaries.
* **OUTPUTEVENTS**: Output information for patients while in the ICU
* **PROCEDUREEVENTS\_MV**: Patient procedures for the subset of patients who were monitored in the ICU using the iMDSoft MetaVision system.

The following tables contain data collected in the hospital record system:

* **CPTEVENTS**: Procedures recorded as Current Procedural Terminology (CPT) codes
* **DIAGNOSES\_ICD**: Hospital assigned diagnoses, coded using the International Statistical Classification of Diseases and Related Health Problems (ICD) system
* **DRGCODES**: Diagnosis Related Groups (DRG), which are used by the hospital for billing purposes.
* **LABEVENTS**: Laboratory measurements for patients both within the hospital and in out patient clinics
* **MICROBIOLOGYEVENTS**: Microbiology measurements and sensitivities from the hospital database
* **PRESCRIPTIONS**: Medications ordered, and not necessarily administered, for a given patient
* **PROCEDURES\_ICD**: Patient procedures, coded using the International Statistical Classification of Diseases and Related Health Problems (ICD) system

The following tables are dictionaries:

* **D\_CPT**: High-level dictionary of Current Procedural Terminology (CPT) codes
* **D\_ICD\_DIAGNOSES**: Dictionary of International Statistical Classification of Diseases and Related Health Problems (ICD) codes relating to diagnoses
* **D\_ICD\_PROCEDURES**: Dictionary of International Statistical Classification of Diseases and Related Health Problems (ICD) codes relating to procedures
* **D\_ITEMS**: Dictionary of ITEMIDs appearing in the MIMIC database, except those that relate to laboratory tests
* **D\_LABITEMS**: Dictionary of ITEMIDs in the laboratory database that relate to laboratory tests

Derived tables

The MIMIC-II database contained a variety of derived tables which simplified use of the database. For example, a commonly used table was the ICUSTAY\_DETAIL table, which provided additional information summarizing a patient’s ICU stay. The database also contained derived parameters commonly required by studies, such as severity scores. In MIMIC-III, we have made a conscious decision to *not* include any derived tables or calculated parameters as far as is possible. Instead, we encourage the community to produce and share scripts which can be run to create these tables or parameters. This has many advantages: it keeps the distinction between raw data and calculated data, it encourages users to validate the scripts which derive the data, and allows for as many scripts as is conceivable without cluttering the database for all users. We have provided a set of scripts at the mimic-code repository, which can be found here:

<http://github.com/MIT-lcp/mimic-code>

We will continue to update this repository both with code which we produce as well as with code produced by the community. We encourage users to make pull requests (a feature of git which allows us to integrate community created code) or raise issues regarding code found in the repository. The creation of an active international community building openly available code for capturing a variety of concepts will increase the speed of research on MIMIC-III exponentially - we hope you take the time to investigate the mimic-code repository for anything which may be of use to you, and further contribute any work of your own!