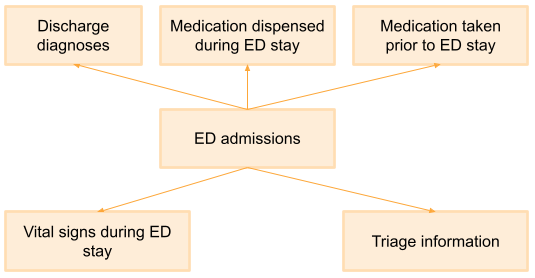
### Overview

The MIMIC-IV-ED (MIMIC-ED) (<https://physionet.org/content/mimic-iv-ed/1.0/>) dataset is a module of the MIMIC-IV (<https://physionet.org/content/mimiciv/0.4/>) dataset, the latter of which is published in August 2020 and contains clinical data on hospital stays sourced from hospital information systems. MIMIC-ED, published in June 2021, contains emergency department (ED) admissions at the Beth Israel Deaconess Medical Center (Boston, MA, USA) between 2011 and 2019. At the time of writing, version 1.0 of MIMIC-ED contains de-identified data from 448,972 ED stays. The dataset is made available freely via PhysioNet ([https://physionet.org/](https://physionet.org/content/mimic-iv-ed/1.0/)), a repository for medical research data, and is intended for both research and educational purposes.

#### Data

MIMIC-ED follows a star-like structure (Figure 1) around the edstays table, which contains two identifiers through which the other tables are linked (subject\_id referring to a patient and stay\_id referring to an ED stay of a patient). The other five tables contain information that was documented during a patient's stay at the ED: discharge diagnoses (diagnosis), medication taken prior to the ED stay (medrecon), medication dispensed during the ED stay (pyxis), information collected at the time of triage (triage), and aperiodic vital signs measured during the ED stay (vitalsign). A description of each table and its elements, including background, license, access, and citation information is provided through the MIMIC-ED page at PhysioNet.

* DOI: <https://doi.org/10.13026/77z6-9w59>
* Data (ZIP file, when access is provided): <https://physionet.org/content/mimic-iv-ed/get-zip/1.0/>



**Figure 1.** MIMIC-ED table structure.

#### **Objective and scenario**

MIMIC-ED can be used for educational and research purposes. PhysioNet distributes the data to its credentialed users that, after signing a data use agreement, can download the raw data (CSV) and/or access the data through services such as the Google Cloud Platform. The objective of this exercise is to create a FHIR version of MIMIC-ED and assess the data FAIRness before and after implementing FHIR. For the FAIRness assessment, we use the [RDA FAIR Data Maturity Indicators](https://confluence.hl7.org/x/rJw7Bg). By doing so, we can 1) contribute to the MIMIC-ED community by creating a FHIR model of the data and providing an exemplar implementation of a FHIR server that can serve MIMIC-ED data, and 2) showing what parts of FAIR can be addressed in this particular case when using FHIR. The lessons learned can be used for other cases where the implementation of the FAIR principles using FHIR is desirable.

### Initial assessment of FAIRness using the RDA FAIR indicators

We assessed the FAIRness of MIMIC-ED before the implementation of FHIR. In other words, the FAIRness of MIMIC-ED as it is being distributed by PhysioNet. The full assessment is available [here (Google Drive)](https://docs.google.com/spreadsheets/d/1N0SKq7iT-MOqCxl3rZAYQsUBGbrh9s5QDHazetFzrGA/edit?usp=sharing). In summary, MIMIC-ED passed most (except one) indicators for Findability and Accessibility. The one indicator that it did not pass is the requirement for metadata to be available when the data is no longer available, we can assume this is true but it has not been specified by PhysioNet. When we look at Interoperability and Reusability we can see a clear use case for using FHIR. MIMIC-ED does not use any FAIR-compliant vocabularies for annotating its data semantics, nor does it comply with a standard data model. Hence, implementing FHIR would probably contribute to the Interoperability and Reusability aspects of FAIR in the case of MIMIC-ED.

**Table 1.**FAIRness assessment of MIMIC-ED before implementing FHIR.

|  | **Findable** | **Accessible** | **Interoperable** | **Reusable** |
| --- | --- | --- | --- | --- |
| **Passed RDA indicators** | 7 / 7 Fully implemented | 11 / 12 Fully implemented | 3 / 12 Not/partly implemented | 2 / 10 Not/partly implemented |
| **Positives** | * (Meta)data are findable through a globally unique DOI * Metadata includes data access instructions and license information * Metadata is findable through a search engine | * Clear data access instructions * (Meta)data can be accessed manually * Resolvable identifiers for (meta)data * HTTP/HTTP GET for (meta)data access * Authentication and authorization through PhysioNet | * Data available as CSV files * Metadata refers to and qualifies other data | * Metadata contains license information * Data are expressed in compliance with a machine- understandable standard (CSV) |
| **Negatives** |  | * No guarantee or mention that metadata will stay available when the data is no longer available | * (Meta)data don't use FAIR-compliant vocabularies for annotating data semantics * (Meta)data don't use a standardised format for  knowledge representation | * No standard or machine-understandable reuse license (PhysioNet license, free-text) * Metadata doesn't include provenance information * (Meta)data don't comply with a community standard (ongoing efforts to use OMOP CDM) |

### **FHIR implementation**

By using FHIR we want to improve the machine-readability/interoperability/FAIRness of the MIMIC-ED dataset. We took the following steps.

* Model the data and metadata conforming the FHIR data model (using the Patient, Encounter, Condition, MedicationStatement, MedicationDispense, Observation, and Procedure resources)
* Set up a FHIR facade server that can serve the (meta)data via the FHIR REST API
* Enable SEARCH operations that allow queries such as “retrieve all ED stays of patient X” or “which patients were discharged from the ED with diagnose Y”



**Figure 2.** Exemplar implementation of a FHIR facade for serving MIMIC-ED data.