

## MIMIC IV Datenqualitätsregeln

### *Synthetisches Dokument im Rahmen des MI-Lab*

Titel	Vollständigkeit (Completeness)
Definition	Ein Datensatz muss alle notwendigen Attribute enthalten. Attribute müssen alle notwendigen Daten enthalten.
Referenz	Mapping gegen Daten, die alle vollständigen Attribute aufweisen.
Regel	<p>Das Attribut „subject_id“ identifiziert einen Patienten und muss immer gefüllt sein.</p> <p>Ein vollständiger Patientendatensatz besteht aus den folgenden Attributen:</p> <ul style="list-style-type: none"><li>- subject_id</li><li>- gender</li><li>- anchor_age</li><li>- anchor_year</li><li>- anchor_year_group</li><li>- dod</li></ul> <p>Das Attribut „hadm_id“ identifiziert eine Hospital Admission und muss immer gefüllt sein mit einer zugehörigen Patienten „subject_id“ Attribut.</p> <p>Ein vollständiger Hospital Admission Datensatz besteht aus den folgenden Attributen:</p> <ul style="list-style-type: none"><li>- subject_id</li><li>- hadm_id</li><li>- admittance</li><li>- dischtime</li><li>- admission_type</li></ul>

Titel	Genauigkeit (Accuracy)
Definition	Die Daten müssen in der jeweils geforderten Exaktheit vorliegen.
Referenz	Vergleich der Formate der Zahlenwerte zu den definierten Regeln.
Regel	<p>Zahlenwerte sind entweder als ganze Zahl (##) oder mit einer Stelle hinter dem Komma (##.#) anzugeben.</p> <p>Ein Datum ist immer in dem Format (yyyy-MM-dd hh:mm:ss) anzugeben.</p>

Titel	Redundanzfreiheit (Nonredundant)
Definition	Innerhalb der Datensätze dürfen keine Dubletten vorkommen.
Referenz	Datenelement, das gegen sich selbst oder sein Gegenstück in einem anderen Datensatz oder einer Datenbank verglichen wurde.
Regel	<p>Jedes Patienten Attribut „subject_id“ darf nur einmal vorkommen.</p> <p>Ausgenommen sind:</p> <ul style="list-style-type: none"><li>- icd_code</li><li>- drg_type, drg_code</li><li>- labevent_id</li><li>- transfer_id</li><li>- services</li><li>- procedures</li><li>- prescriptions</li></ul>