



























Diccionario de datos

-  ADMISSIONS
-  CALLOUT
-  CAREGIVERS
-  CHARTEVENTS
-  CPTEVENTS
-  D_CPT
-  D_ICD_DIAGNOSES
-  D_ICD_PROCEDURES
-  D_ITEMS
-  D_LABITEMS
-  DATETIMEEVENTS
-  DIAGNOSES_ICD
-  DRGCODES
-  ICUSTAYS
-  INPUTEVENTS_CV
-  INPUTEVENTS_MV
-  LABEVENTS
-  MICROBIOLOGYEVENTS
-  NOTEEVENTS
-  OUTPUTEVENTS
-  PATIENTS
-  PRESCRIPTIONS
-  PROCEDUREEVENTS_MV
-  PROCEDURES_ICD
-  SERVICES
-  TRANSFERS

Descripción general

- La base de datos MIMIC-3 consta de 26 tablas.
- Cada tabla contiene el registro de cada paciente (en cada fila) con un campo específico (columnas)
- Las tablas que comienzan con 'D_' son diccionarios y proporcionan definiciones para identificadores.
- “_MV” y “_CV” en los nombres de las tablas representan diferentes sistemas de información utilizados para recopilar datos.
- CV: Philips Carevue, 2001-2008
- MV: Metavisión de iMDSOft, 2008-2012

Información del paciente

- Cada paciente es único con su propio "subject_id".
- Cada ingreso hospitalario de un paciente es único con "hadm_id".
- Cada estancia en la UCI de un paciente es única con "icustay_id".
- Un subject_id se puede asociar con múltiples hadm_ids cuando un paciente tuvo múltiples admisiones.

SUBJECT_ID	HADM_ID	ICUSTAY_ID
22067	100566	258026
22067	100566	253296
22067	100566	232637

- Un hadm_id se puede vincular a múltiples icustay_id cuando un paciente tuvo múltiples estancias en la UCI durante una admisión. (p. ej., transferencia entre varias UCI)

PACIENTES

- Sexo, fecha de nacimiento (fecha de nacimiento).
- DOD (fecha de muerte), indicador de caducidad (ya sea que un paciente haya muerto o no)
- La edad de cierto paciente de un punto de tiempo en el registro se puede calcular restando un cierto tiempo de registro - dob.
- Tenga en cuenta que todas las fechas en la base de datos se cambian aleatoriamente con el proceso de desidentificación. Pero es consistente a lo largo de los registros de un paciente.

ADMISIONES

- Tiempo de admisión y alta, tiempo de muerte si murió en la admisión.
- Lugar de ingreso y alta en el hospital.
- Seguro, idioma, etnia y estado civil.
- Diagnóstico, ** pero este diagnóstico generalmente lo asigna el médico de admisión y no usar una ontología sistémica (como el código ICD9).
- El diagnóstico final se puede encontrar en la tabla DIAGNOSIS_ICD.

ICUSTAYS

- Hora de entrada/salida
- LOS (duración de la estancia), los valores se normalizan a 1,0 = 24 horas

MEDICAMENTO

Tres tablas contienen información sobre la medicación:

1) RECETAS

2) INPUTEVENTS_CV

3) INPUTEVENTS_MV

PRESCRIPCIONES

- Relacionado con la medicación orden entradas.
- Nombres de medicamentos, código NDC (versión de 11 dígitos)
- El Código Nacional de Drogas (NDC) es originalmente un identificador numérico de 3 segmentos y 10 dígitos.
- Los Centros de Servicios de Medicare y Medicaid (CMS) han creado un derivado NDC de 11 dígitos. Este formato ha sido seleccionado por la regulación HIPAA, por lo que la base de datos de otras agencias gubernamentales como UMLS (que contiene ontologías RXNORM, NDF-RT) está utilizando este NDC de 11 dígitos.

DRUG	DRUG_NAME_POE	DRUG_NAME_GENERIC	FORMULARY_DRUG_CD	GSN	NDC
Ipratropium Bromide MDI	Ipratropium Bromide MDI	Ipratropium Bromide HFA	IPRAPF	059081	00597008717
Albuterol Inhaler	Albuterol Inhaler	Albuterol Inhaler (ProAir)	PROAIR8.5H	028090	59310057920
Heparin	Heparin	Heparin Sodium	HEPA5I	006549	63323026201

INPUTEVENTS

- Las ENTRADAS son cualquier líquido que ha sido administrado al paciente como alimentación oral o por sonda o soluciones intravenosas que contienen medicamentos.
- Las entradas existen en dos tablas separadas:
 - INPUTEVENTS_CV contiene entradas de CareVue
 - INPUTEVENTS_MV contiene entradas de Metavision

ENTRADAS CV

- Para los datos de CareVue, solo CHARTTIME está disponible.
- Las columnas TASA y CANTIDAD son asincrónicas, y originalmente almacenadas en diferentes tablas. (En MIMIC-3, estos están en una tabla).
- Volúmen es de entrada (p. ej., 50 ml de solución salina normal) normalmente se registraría cada hora (aunque a veces el período era más largo).
- TASA del fármaco sólo se actualizaba cuando se realizaba un cambio o verificación de la tarifa por parte del personal clínico.

CHARTTIME	VOLUME	VOLUMEUOM	CHARTTIME	RATE	RATEUOM
			09:00	1	mL/min
10:00	60	mL			
11:00	60	mL			
			11:30	0.5	mL/min
12:00	45	mL			

- Suponiendo que se trata de registros de eventos de entrada de un paciente de un determinado medicamento.
- El volumen se registra solo cada hora.
- No hay hora de inicio disponible. (CHARTTIME del volumen es la hora de finalización)
- Sin embargo, es razonable suponer que la medida del volumen corresponde a una hora.
- A continuación, podemos ver que la tasa se tituló a 0.5
- Para el período de 11:00 a 12:00 hubo media hora de entrega a 1 mL/min y media hora de entrega a 0,5 mL/min.
- =>Dando como resultado un volumen total de 45 ml administrados durante una hora entre las 11:00 y las 12:00.
- Recuerda eso VOLUMEN normalmente sería registrado cada hora.
- TASA es solo se actualiza cuando hay un cambio o verificación de la tasa fue realizada por el personal clínico.

- Para tarifas, CHARTTIME corresponderá a una hora de inicio (cuando el fármaco se fijó a esa tasa).
- Para volúmenes, CHARTTIME corresponderá a una hora de finalización.

SOLICITAR ID

- Vincula varios elementos contenidos en la misma solución.
- Por ejemplo, cuando se administra un fármaco de noradrenalina y una solución de NaCl.

HORA DEL CARTA	ARTÍCULO	SOLICITAR ID
10:00	noradrenalina (droga)	201
10:00	NaCl (solución)	201

- Tanto la noradrenalina como el NaCl se encuentran en filas distintas, pero tienen lo mismo SOLICITAR ID.

LINKORDERID

- Vincula el mismo pedido a través de múltiples instancias.
- Por ejemplo, cuando la tasa de entrega para la solución con noradrenalina y NaCl es cambiada.
- Dos nuevas filas que comparten el mismo nuevo SOLICITAR ID se generarán, pero el LINKORDERID será lo mismo.

HORA DEL CARTA	ARTÍCULO	SOLICITAR ID	LINKORDERID	TASA
10:00	noradrenalina	201	5001	1.0
10:00	NaCl	201	5001	1.0

11:00	noradrenalina	202	5001	0.5
11:00	NaCl	202	5001	0.5

- Para los datos de Metavision, no existe el concepto de un volumen en la base de datos, solo una TASA.
- Todas las entradas se registran con una HORA DE INICIO y una HORA DE FINALIZACIÓN.
- Como resultado, los volúmenes en la base de datos para los pacientes de Metavision son derivado de las tarifas.
- Además, los tiempos exactos de inicio y finalización de los medicamentos son fácilmente deducibles.
- La columna ORDERID sería idéntica para cada instancia de NaCl y noradrenalina que correspondiera a la misma solución/ velocidad.
- Es decir, para la infusión administrada entre las 18:20 y las 18:25, tanto el NaCl como la noreadrenalina tendrían el mismo ORDERID.
- LINKORDERID vincularía aún más el medicamento en todas las administraciones, incluso cuando se cambia la tasa.

PRUEBAS DE LABORATORIO

Tres tablas contienen registro de prueba de laboratorio:

- CARTAEVENTOS
- LABEVENTOS
- EVENTOS DE MICROBIOLOGIA

Dos tablas contienen información de la etiqueta itemid

- D_ITEMS (para CHARTEVENTS y MICROBIOLOGYEVENTS)
- D_LABITEMS (para LABEVENTS)

CHARTEVENTS

- La tabla CHARTEVENTS contiene todos los datos graficados disponibles para los pacientes.
- Contiene signos vitales de rutina y cualquier información adicional relevante para su atención: configuración del ventilador, valores de laboratorio, estado del código, estado mental, etc.

- Como resultado, CHARTEVENTS contiene la mayor parte de la información sobre la estadía de un paciente más que las pruebas de laboratorio.
- Además, aunque los valores de laboratorio se capturan en otro lugar (LABEVENTS), se repiten con frecuencia dentro de CHARTEVENTS.
- En los casos en que exista discrepancia entre las mediciones, se debe tomar la tabla LABEVENTS como la verdad fundamental.

Itemid

- Identificador de un tipo de medida en la base de datos
- La etiqueta y la definición de itemid se pueden encontrar en la tabla D_ITEMS
- Por ejemplo, itemid 212 en CHARTEVENTS es 'insuficiencia cardíaca'
- Valor: el valor medido para el artículo.
- Si este valor es numérico, VALUENUM contiene los mismos datos en formato numérico.

LABEVENTOS

- Esta tabla contiene información sobre mediciones de laboratorio.
- La etiqueta de cada itemid se puede encontrar en la tabla 'D_LABITEMS'.
- El proceso de recopilación de datos es el siguiente
 - 1) El personal clínico adquiere un líquido de un sitio en el cuerpo del paciente
 - 2) El fluido tiene un código de barras asociado con el paciente y una marca de tiempo
 - 3) Se ejecuta el análisis de laboratorio y se devuelve el resultado dentro de 4 a 12 horas.
- La columna charttime registra cuándo se crea una observación.
- Es el proxy más cercano al momento en que se midieron realmente los datos.

HADM_ID	ITEMID	CHARTTIME	VALUE	VALUENUM	VALUEUOM	FLAG
NULL	51006	2101-10-04 15:24:00	42	42	mg/dL	abnormal
NULL	51137	2101-10-04 15:24:00	NORMAL	NULL	NULL	NULL
NULL	51146	2101-10-04 15:24:00	0.5	0.5	%	NULL

LABEVENTOS (cont.)

- El campo de valor contiene la medida de valor para el elemento de laboratorio.
- El campo BANDERA indica si el valor se considera anormal o no utilizando umbrales predefinidos.

- Un paciente sin ID de ingreso hospitalario (hadm_id) es un valor de laboratorio obtenido como paciente ambulatorio.

D LABITEMS

- Contiene identificadores asociados con mediciones de laboratorio.
- Cada entrada tiene códigos LOINC.
- Fluido: la sustancia sobre la cual se realizó la medición.
- Categoría: información de nivel superior en cuanto al tipo de medición.

	ITEMID	LABEL	FLUID	CATEGORY	LOINC_CODE
1	50800	SPECIMEN TYPE	BLOOD	BLOOD GAS	NULL
2	50801	Alveolar-arterial Gradient	Blood	Blood Gas	19991-9
3	50802	Base Excess	Blood	Blood Gas	11555-0

EVENTOS DE MICROBIOLOGÍA

- Contiene información microbiológica, incluidas las pruebas realizadas y las sensibilidades.
- CHARTTIME: hora en la que se cartografió una observación y, por lo general, es el indicador más cercano a la hora en que se midieron realmente los datos.
- SPEC_ITEMID y SPEC_TYPE_DESC: información sobre la muestra.
- ORG_ITEMID y ORG_NAME: información sobre el organismo.
- AB_ITEMID y AB_NAME: información sobre anticuerpos.
- INTERPRETACIÓN: “S” es sensible, “R” es resistente, “I” es intermedia y “P” es pendiente.

CHARTTIME	SPEC_ITEMID	SPEC_TYPE_DESC	ORG_ITEMID	ORG_NAME	ISOLATE_NUM	AB_ITEMID	AB_NAME
2165-06-03 08:28:00	70079	URINE	80004	KLEBSIELLA PNEUMONIAE	1	90022	AMPICILLIN/SULBACTAM
2165-06-03 08:28:00	70079	URINE	80004	KLEBSIELLA PNEUMONIAE	1	90019	CIPROFLOXACIN
2165-06-03 08:28:00	70079	URINE	80004	KLEBSIELLA PNEUMONIAE	1	90018	CEFTRIAXONE
2165-06-03 08:28:00	70079	URINE	80004	KLEBSIELLA PNEUMONIAE	1	90017	CEFTAZIDIME
2165-06-03 08:28:00	70079	URINE	80004	KLEBSIELLA PNEUMONIAE	1	90013	TOBRAMYCIN

- Los productos de desecho son fluidos que han sido excretados por el paciente, como la producción de orina, o extraídos del paciente, por ejemplo, a través de un drenaje.
- Todos los registros de salidas están en la tabla OUTPUTEVENTS.
- La hora a la que se mide la salida se registra en la columna CHARTTIME.
- No hay una hora de inicio registrada con salidas - CHARTTIME corresponde a la hora en que se emitió el volumen.
- El volumen de salida se registra en la columna VALOR y la unidad de medida se proporciona en la columna VALUEUOM (generalmente mililitros o mL).

- Por lo general, es razonable suponer que cualquier salida registrada corresponde al intervalo entre CHARTTIME actual y CHARTTIME anterior para el mismo elemento.

DIAGNOSTICO

- La tabla DIAGNOSES_ICD contiene tuplas de identificación del paciente (subject_id), identificación de admisión (hadm_id) y código ICD-9.
- No hay información de marca de tiempo disponible.
- El campo de código para los códigos de diagnóstico principal y otros códigos ICD-9-CM tiene una longitud de seis caracteres, con el punto decimal implícito entre el tercer y el cuarto dígito para todos los códigos de diagnóstico que no sean los códigos V.
- El decimal está implícito para los códigos V entre el segundo y el tercer dígito.
- Número de secuencia: los diagnósticos de ICD se ordenan por prioridad, y el orden tiene un impacto en el reembolso del tratamiento.
- La tabla D_ICD_DIAGNOSES contiene una etiqueta para cada código ICD-9.

PROCEDIMIENTOS

- La tabla PROCEDURES_ICD contiene un registro de procedimiento.
- Estructura similar a la tabla DIAGNOSES_ICD:
 - Identificación del paciente, identificación de admisión, código de procedimiento ICD-9 y número de secuencia.
 - La tabla D_ICD_PROCEDURES contiene una etiqueta para los códigos de procedimiento ICD-9.
 - En esta tabla se utilizan 3882 tipos diferentes de procedimientos.
 - La tabla PROCEDUREEVENTS_MV también contiene registros de procedimientos.
 - Procedimientos de pacientes para el subconjunto de pacientes que fueron monitoreados en la UCI usando el sistema iMDSoft MetaVision.
 - En esta tabla se utilizan 125 tipos diferentes de procedimientos.
 - La tabla CPTEVEENTS contiene códigos de terminología procesal actual (CPT) que facilitan el propósito de facturación.
 - CPT_CD: código CPT original
 - CPT_NUMBER: versión numérica de CPT_CD.
 - SECTIONHEADER: una categoría del código CPT.
 - La tabla D_CPT contiene una etiqueta para el código CPT.

NOTAS CLINICA

- La tabla NOTEEVENTS contiene una nota clínica de texto libre.
- Incluyendo notas médicas y de enfermería, informes de ECG, informes de radiología y resúmenes de alta.
- Categoría completa:

- Gestión de casos
- Consultar
- Resumen de alta electrocardiograma
- Eco (ecoencefalograma)
- General
- Enfermería
- Enfermería/otro
- Nutrición
- Farmacia
- Médico
- Radiología
- Servicios de rehabilitación
- Respiratorio
- Trabajo Social

OTRAS

- Hay algunas tablas que no mencioné en estas diapositivas.
- Un breve es:
 - LLAMADA: Información sobre cuándo se autorizó el alta de la UCI a un paciente y cuándo se le dio de alta realmente.
 - CUIDADORES: Todo cuidador que tenga datos registrados en la base de datos (define CGID).
 - DATETIMEEVENTS: todas las observaciones registradas que son fechas, por ejemplo, hora de diálisis o inserción de líneas.
 - DRGCODES: Grupos Relacionados con el Diagnóstico (GRD), que son utilizados por el hospital para fines de facturación.
 - SERVICIOS: El servicio clínico bajo el cual está registrado un paciente.
 - TRASLADOS: Movimiento de pacientes de cama a cama dentro del hospital, incluyendo ingreso y alta en UCI.