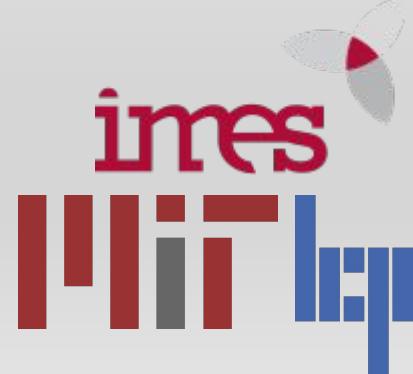


The eICU Collaborative Research Database

Alistair Johnson - aewj@mit.edu

Laboratory for Computational Physiology
Institute for Medical Engineering & Science
Massachusetts Institute of Technology



Outline

- Brief overview of telehealth and tele ICUs
- The eICU Collaborative Research Database
- Deep dive into the tables
- (Optional!) Install the demo on your local computer

What is a TeleICU?

TELEICU DEFINED

TeleICU is a centralized remote model of critical care delivery that uses special tools to leverage clinical expertise across a variety of clinical settings and geographical areas.

STRUCTURE TYPES

Care teams working in a defined physical structure are centralized. When care providers are not confined to a physical space it is considered decentralized.

TELEICU MODELS OF CARE

- Continuous Care is monitoring without interruption
- Responsive (Reactive) Care is unscheduled and prompted by an alert, alarm, page, telephone call...
- Scheduled Care are scheduled periodic consultations



SPECIAL TOOLS

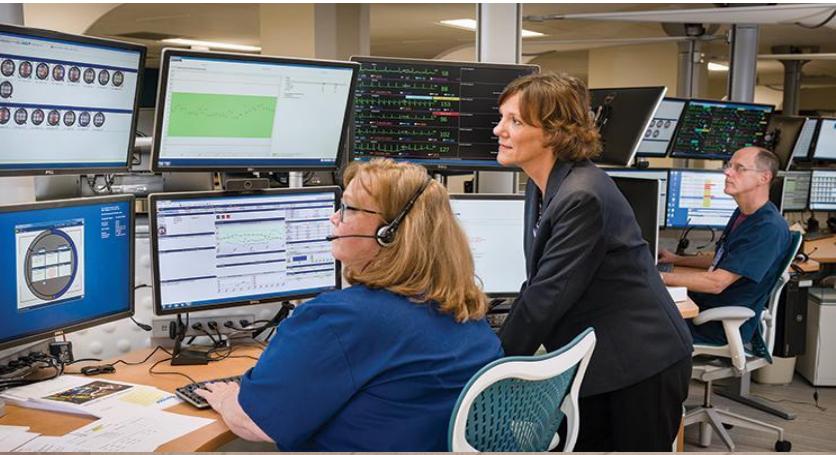
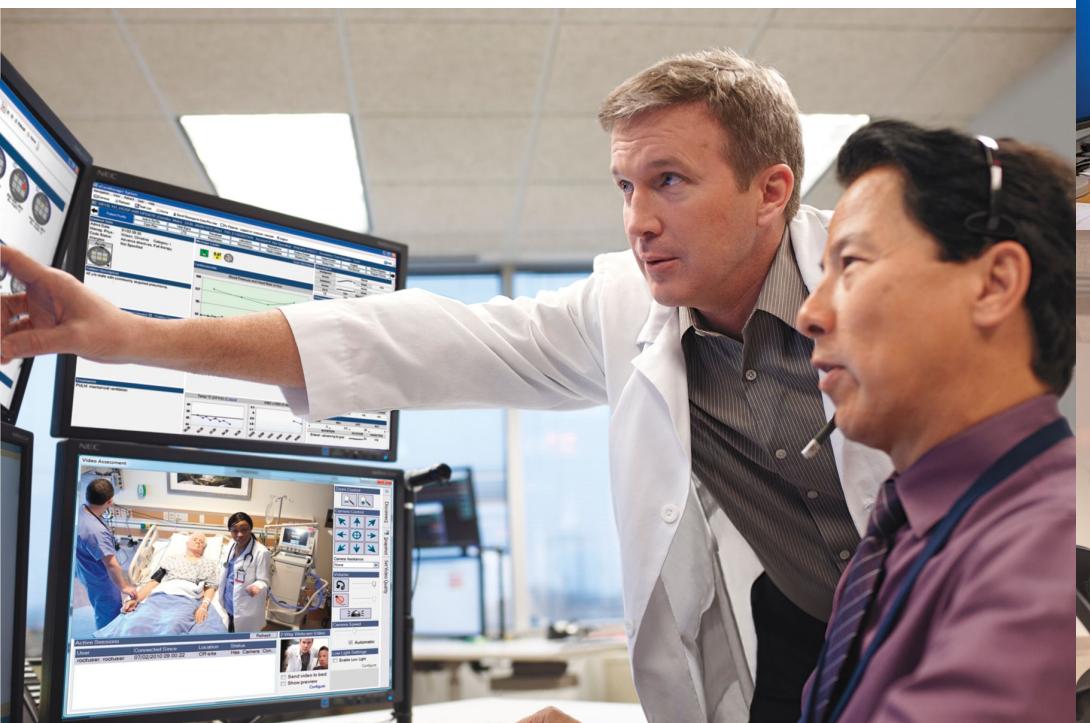
- Alerts are tools designed to recognizing signs of clinical deterioration
- Clinical decision support systems reduce the need to remember everything
- Tools support continuous capture, analyses, and disseminating of information
- Audio-video tools support remote assessment

1. ATA TeleICU Practice Guidelines Work Group. Guidelines for TeleICU Operations. ATA Standards & Guidelines. 2014;2014(June 21). <http://www.americantelemed.org/resources/standards/ata-standards-guidelines#.U7gm1hbZf1p>.

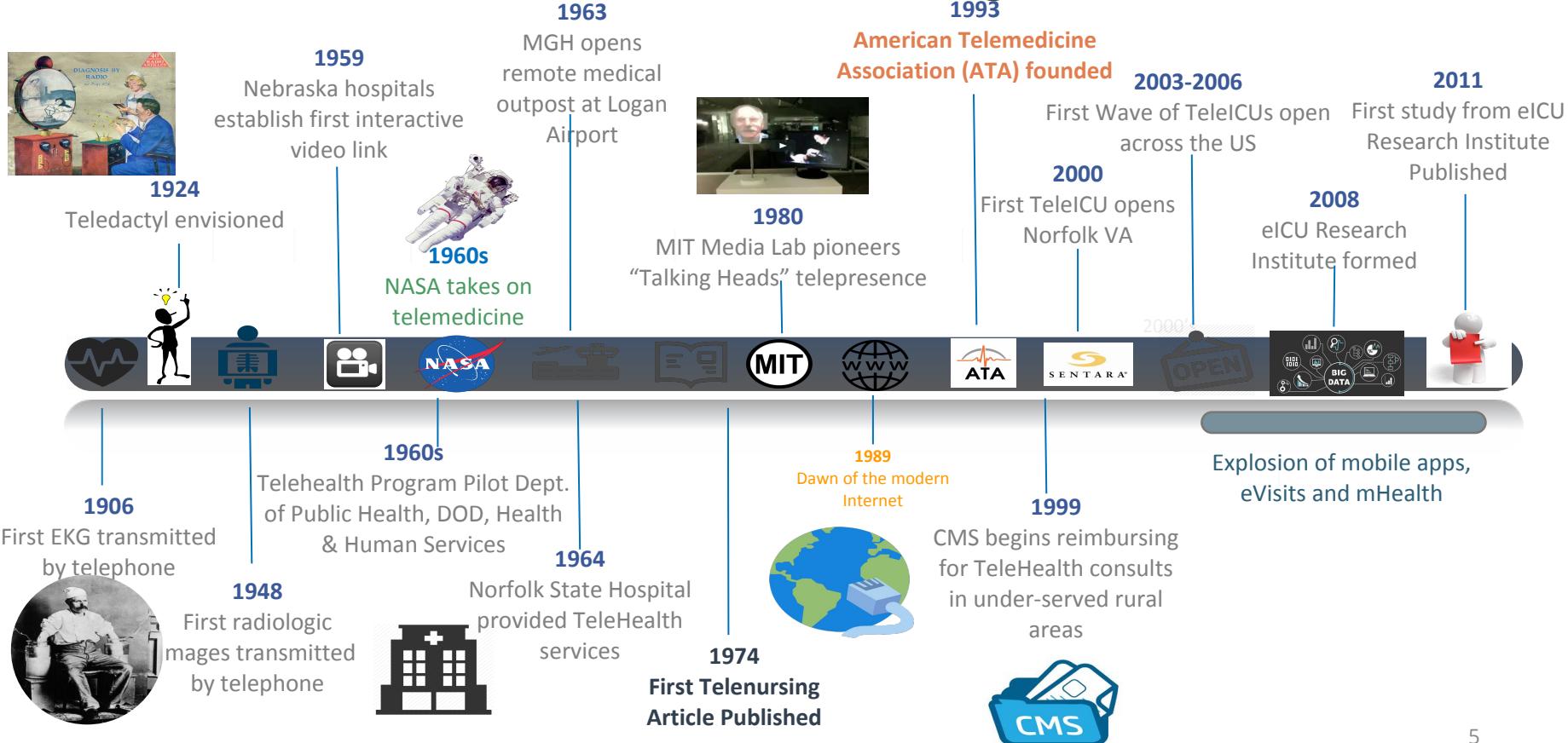
2. Lilly CM, Zubrow MT, Kempner KM, et al. Critical care telemedicine: evolution and state of the art. Crit Care Med. 2014;42(11):2429-2436.

3. Lilly CM, Motzkus C, Rincon T, Cody SE, Landry K, Irwin RS. ICU Telemedicine Program Financial Outcomes. Chest. 2017;151(2):286-297.

What is an eICU?

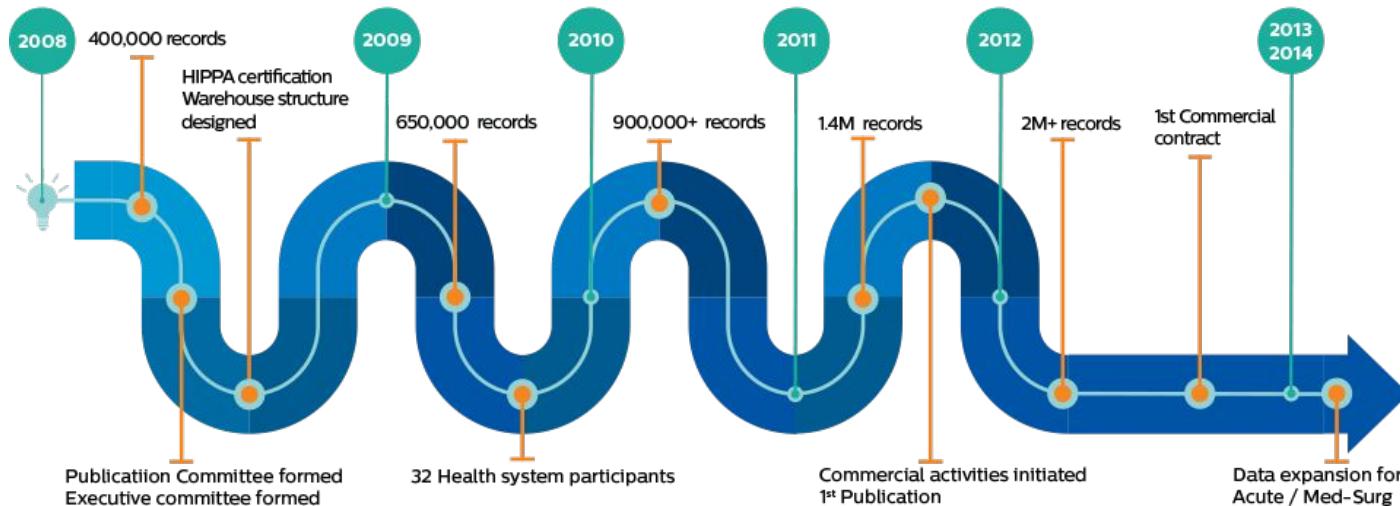


Evolution of Telehealth - founding of eRI

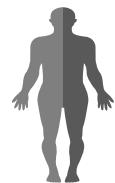


eICU Research Institute - eRI

<https://www.usa.philips.com/healthcare/solutions/enterprise-telehealth/eri>



eRI studies



243,553 eligible patients in 271 ICUs 2008

Lilly C. et al Benchmark data from more than 240,000 adults that reflect the current practice of critical care in the United States. *Chest*. 2011

194,722 eligible patients 2008-2010

Badawi O et al. Association between intensive care unit-acquired dysglycemia and in-hospital mortality*. *Critical Care Medicine*. 2012

207,702 eligible patients in 344 ICUs 2008-2010

Waite MD, et al. Intensive care unit acquired hypernatremia is an independent predictor of increased mortality and length of stay. *Journal of Critical Care*. 2013

2,014 eligible patients in 344 ICUs

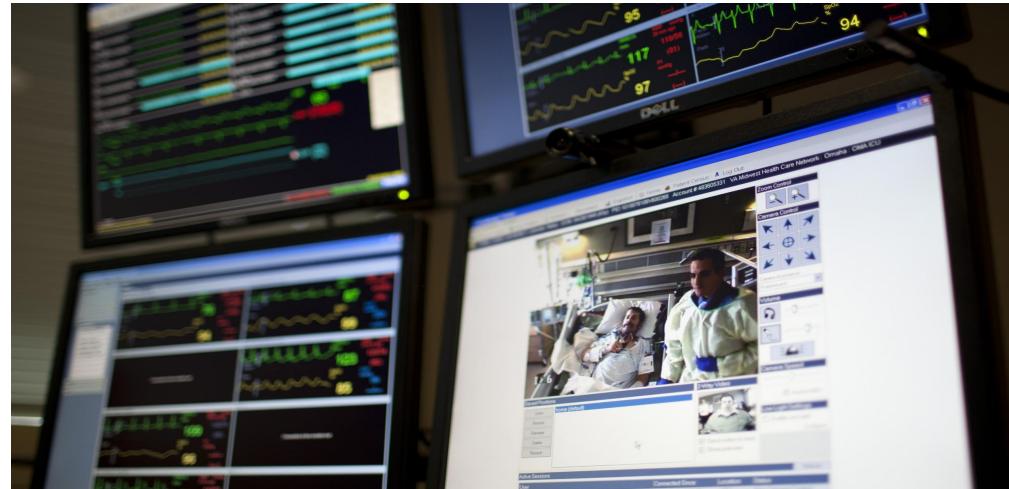
2008-2010

McPhee LC et al. Single-Dose Etomidate Is Not Associated With Increased Mortality in ICU Patients With Sepsis: Analysis of a Large Electronic ICU Database*. *Critical Care Medicine*. 2013

The eICU Collaborative Research Database

- Philips eRI partner with MIT-LCP on research projects
- First research project: release data!

- 200,000+ admissions
- 160,000+ unique patients
- Intensive care units (ICUs) across the US
- Multicenter - 300+ hospitals



<https://eicu-crd.mit.edu>

How did we get this data?

Philips eRI performed the Extract-Transform-Load process

MIT-LCP...

- Wrote an algorithm to match together the same patient (-> *uniquePid*)
- Selected a representative sample of the entire database
- Reformatted the data and loaded it into a database
- Provided QA

Nomenclature

Tables will be represented as *italicized_bolded_text*

- For example, the ***patient*** table is different than a single patient

Columns will be represented as *italicized_text*

- For example, the *uniquePid* column in the ***patient*** table

I will use case for legibility in this slide (e.g. uniquePid), but the PostgreSQL database by default has all column names in lower case

Overview of tables

Patient info

*patient
hospital*

APACHE-IV

*admissionDx
apacheApsVar
apachePatientResult
apachePredVar*

Care plans

*carePlanCareProvider
carePlanEOL
carePlanGeneral
carePlanGoal
carePlanInfectiousDiseases*

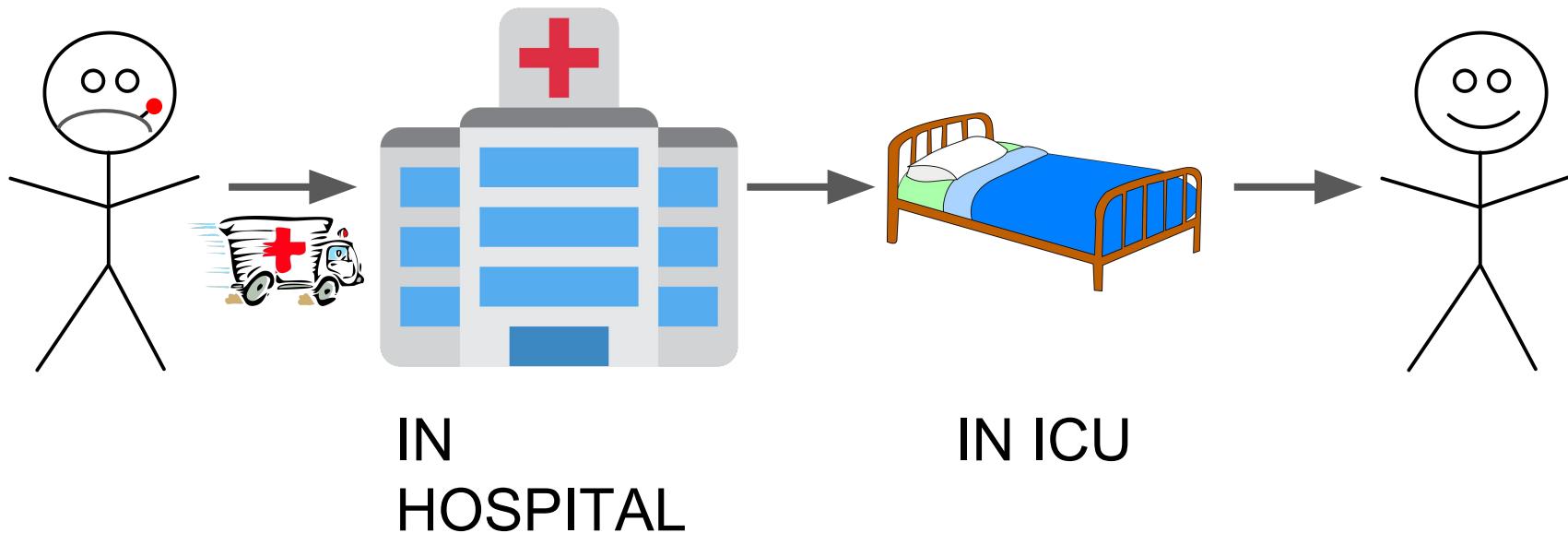
Timestamped data

*admissionDrug
diagnosis
infusionDrug
lab
pastHistory
treatment
medication*

Monitor data

*vitalAperiodic
vitalPeriodic*

Patient's hospital course



Will try to refer back to this image when describing data



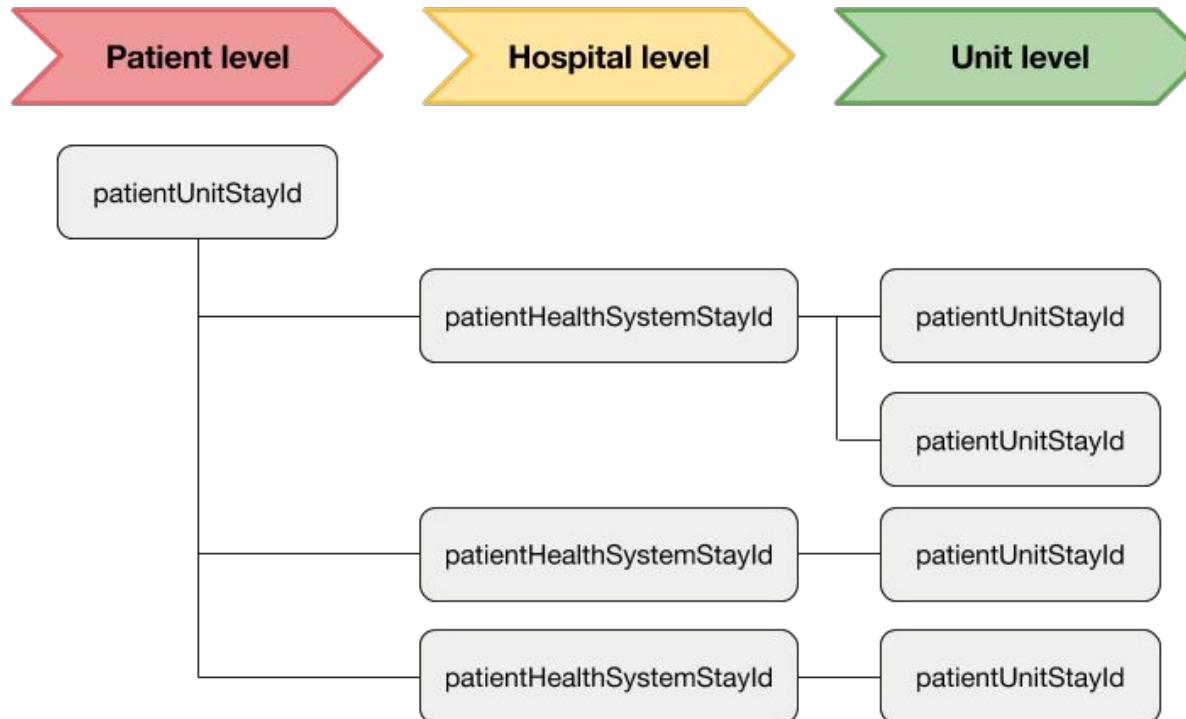
patient, hospital

patient

- *patientUnitStayId* - a unit stay
- *patientHealthSystemStayId* - a hospital stay
- *uniquePID* - a patient stay

patient

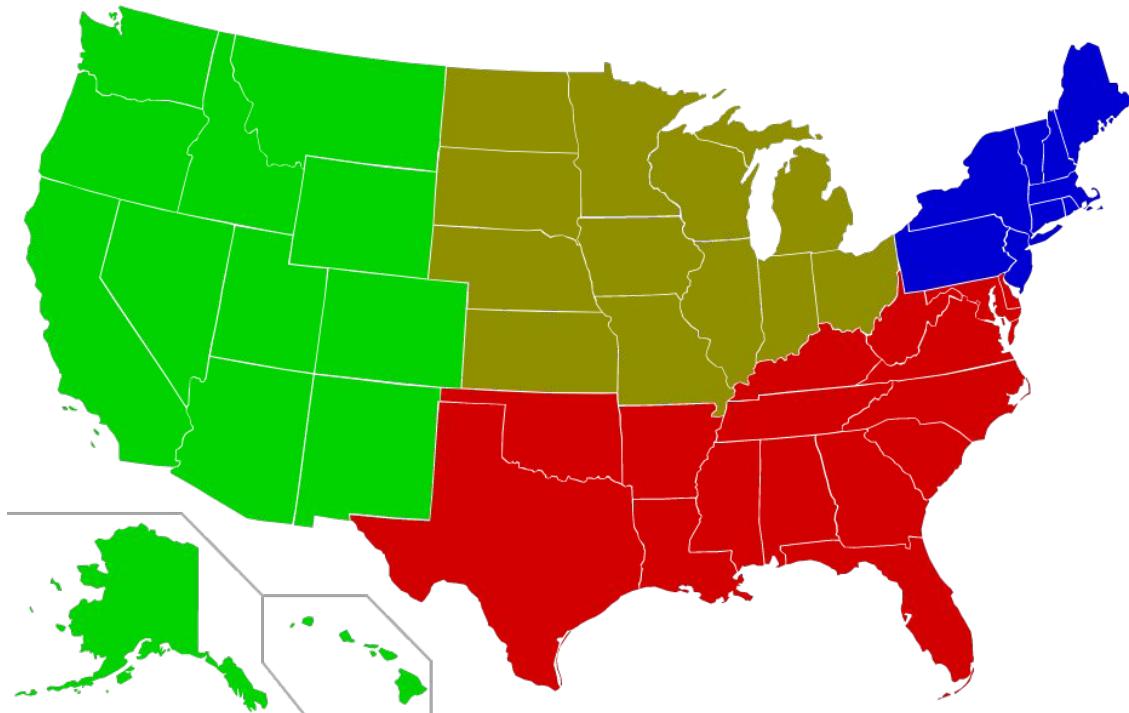
All data in eICU is in one table, called patient, even though it has unique rows for each *ICU* stay



hospital

- Number of beds
- Teaching hospital
- Region
- Results from a survey

region		count
		180
Midwest		106
Northeast		20
West		66
South		87



hospital

- Number of beds
- Teaching hospital
- Region
- Results from a survey

region	count
	180
Midwest	106
Northeast	20
West	66
South	87

teachingstatus	count
t	48
t	24 (5%)
f	387 (85%)

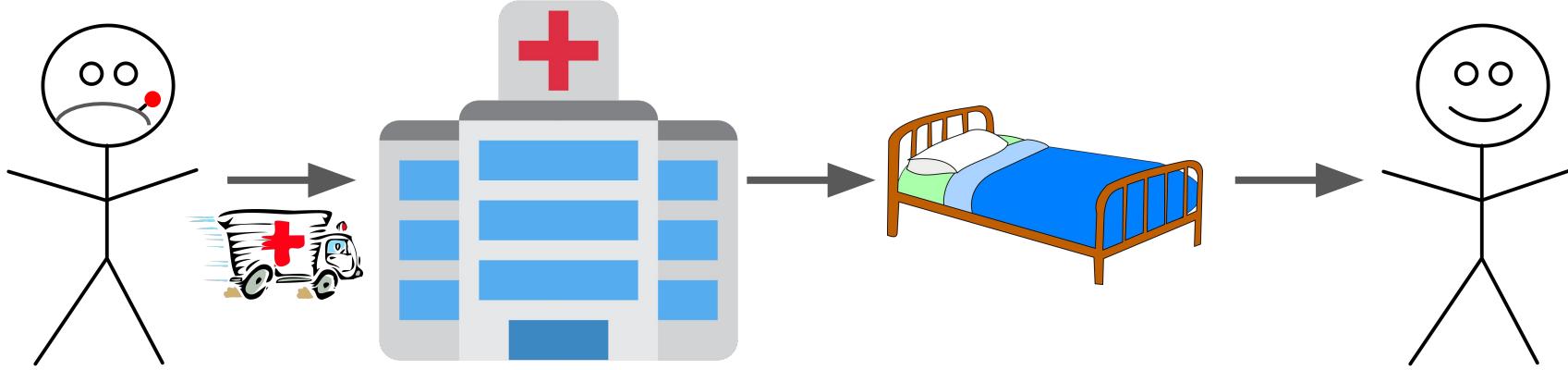
numbedscategory	count
<100	223 (48%)
100 - 249	84 (18%)
250 - 499	81 (17%)
>= 500	46 (10%)
	25 (5%)

Number
of
hospitals
with data

<i>hospitalId</i>	459
<i>region</i>	279
<i>numBedsCategory</i>	239
<i>teachingStatus</i>	411

Not very
many
teaching
hospitals

Patient's hospital course



patient

- *hospitalAdmitOffset*
- *hospitalDischOffset*
- *hospitalId*

hospital

- *numBedsCategory*
- *teachingStatus*
- *region*

patient

- *unitDischOffset*
- *unitType*
- *unitAdmitSource*
- *unitStayType*
- *unitDischargeLocation*
- *wardId*

Overview of tables

Patient info

*patient
hospital*

APACHE-IV

*admissionDx
apacheApsVar
apachePatientResult
apachePredVar*

Care plans

*carePlanCareProvider
carePlanEOL
carePlanGeneral
carePlanGoal
carePlanInfectiousDiseases*

Timestamped data

*admissionDrug
diagnosis
infusionDrug
lab
pastHistory
treatment
medication*

Monitor data

*vitalAperiodic
vitalPeriodic*



APACHE IV

apacheApsVar, apachePredVar, apachePatientResult

Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients*

Zimmerman, Jack E. MD, FCCM; Kramer, Andrew A. PhD; McNair, Douglas S. MD, PhD; Malila, Fern M. RN, MS

Critical Care Medicine: May 2006 - Volume 34 - Issue 5 - pp 1297-1310

doi: 10.1097/01.CCM.0000215112.84523.F0

Feature Articles

- Age
- Acute physiology
- Prior site of healthcare (floor, emergency room)
- Major disease category

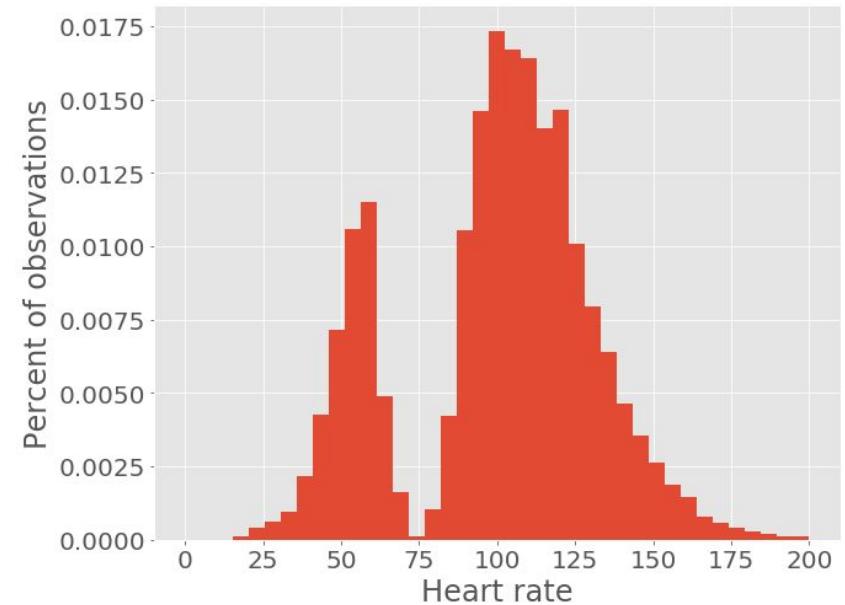
apacheApsVar, apachePredVar, apachePatientResult

- APACHE IV: Excellent mortality prediction model
- Uses concept of “APACHE day” (roughly first 24 hours)
- Collect data -> logistic regression to predict in-hospital mortality
- Exclusion criteria
 - <4 hours LOS
 - Burns
 - Certain transplants
 - In-hospital readmission
 - CABG patients receive special equation with unique covariates

apacheApsVar

- “Worst” Physiology on APACHE day 1
 - Note bimodal distribution - whichever heart rate resulted in the higher score
 - Vitals
 - Labs
 - Some treatments
 - Ventilation, thrombolytic therapy
- Need to interpret these features
according to APACHE data collection
 - For example, heart rate rule appears to be furthest value from 75

Pulse
Mean BP (mmHg)
Temperature (°C)



8	5	0	1	5	7	13	17
≤ 39	40-49	50-99	100-109	110-119	120-139	140-154	≥ 155
23	15	7	6	0	4	7	9
≤ 39	40-59	60-69	70-79	80-99	100-119	120-129	130-139
20	16	13	8	2	0	4	
≤ 32.9	33-33.4	33.5-33.9	34-34.9	35-35.9	36-36.9	≥ 40	

apachePredVar

- Other covariates - combined with APS to create prediction
- Does contain some physiology
- Teaching type
- Admission diagnosis
- Comorbidities
 - *aids, HepaticFailure, Lymphoma, MetastaticCancer, Leukemia, Immunosuppression, Cirrhosis*
- Active treatment -> thrombolytic therapy
- Diabetes, Vent day 1 (overlap with *apacheApsVar* !)

apachePatientResult

- The result of the APACHE IV and APACHE IVa predictions
- Two rows per *patientUnitStayId* !
- Has the APS-III: *acutePhysiologyScore*
- *apacheScore* = APS-III + CHE + Age
- Predictions
 - *predictedICUMortality*, *predictedHospitalMortality*
 - *predictedICULOS*, *predictedHospitalLOS*
 - *predVentDays*
- *apacheScore* - calculated for readmissions
- *predictedHospitalMortality* - won't be calculated for readmissions

admissionDx

- Contains admission diagnosis as codified by APACHE
- Also contains other APACHE information
 - Was the patient admitted from the O.R. or went to the O.R. within 4 hours of admission?
 - Elective?
 - Thrombolytic Therapy received within 24 hours?
- Some information is rare (<2%) but could be useful
 - Pre-op ejection fraction (%)
 - Saphenous vein graft?
 - PTCA done within 24 hours
 - Acute MI location

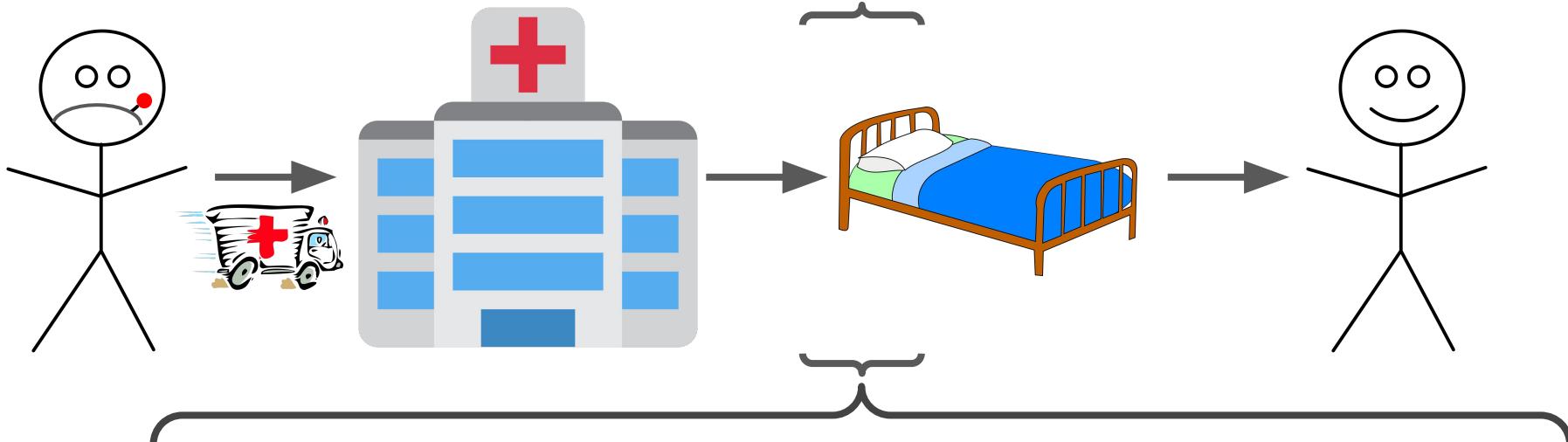
admissionDx

- *admitDxPath*
 - Full description
 - admission diagnosis|All Diagnosis|Non-operative|Diagnosis|Cardiovascular|Rhythm disturbance (atrial, supraventricular)
- *admitDxName*
 - Last part of the path
 - Rhythm disturbance (atrial, supraventricular)
- *admitDxText*
 - The same as admitDxName for diagnoses
 - Rhythm disturbance (atrial, supraventricular)

admissionDx

lvl1	lvl2	numobs
admission diagnosis	All Diagnosis	177863
admission diagnosis	Non-operative Organ Systems	143620
admission diagnosis	Additional APACHE Information	54522
admission diagnosis	Elective	34243
admission diagnosis	Operative Organ Systems	34243

Patient's hospital course



admissionDx

- *admitDxEnteredOffset*
- *admitDxPath*
- *admitDxText*

apacheApsVar

- *Physiology*
- *Some comorbidities*

apachePredVar

- *Comorbidities*
- *Diagnosis*

apachePredVar

- *Comorbidities*
- *Diagnosis*
- *Treatment*

Overview of tables

Patient info

*patient
hospital*

APACHE-IV

*admissionDx
apacheApsVar
apachePatientResult
apachePredVar*

Care plans

*carePlanCareProvider
carePlanEOL
carePlanGeneral
carePlanGoal
carePlanInfectiousDiseases*

Timestamped data

*admissionDrug
diagnosis
infusionDrug
lab
pastHistory
treatment
medication*

Monitor data

*vitalAperiodic
vitalPeriodic*



Timestamped data

pastHistory

- 87.8% of patients have an observation
- Can be used to determine comorbidities
 - insulin dependent diabetes - 12%
 - CHF - 11%
 - No past problems - 6%
- We will use this table to demonstrate how eICU-CRD stores most data

Timestamped data - full path

path

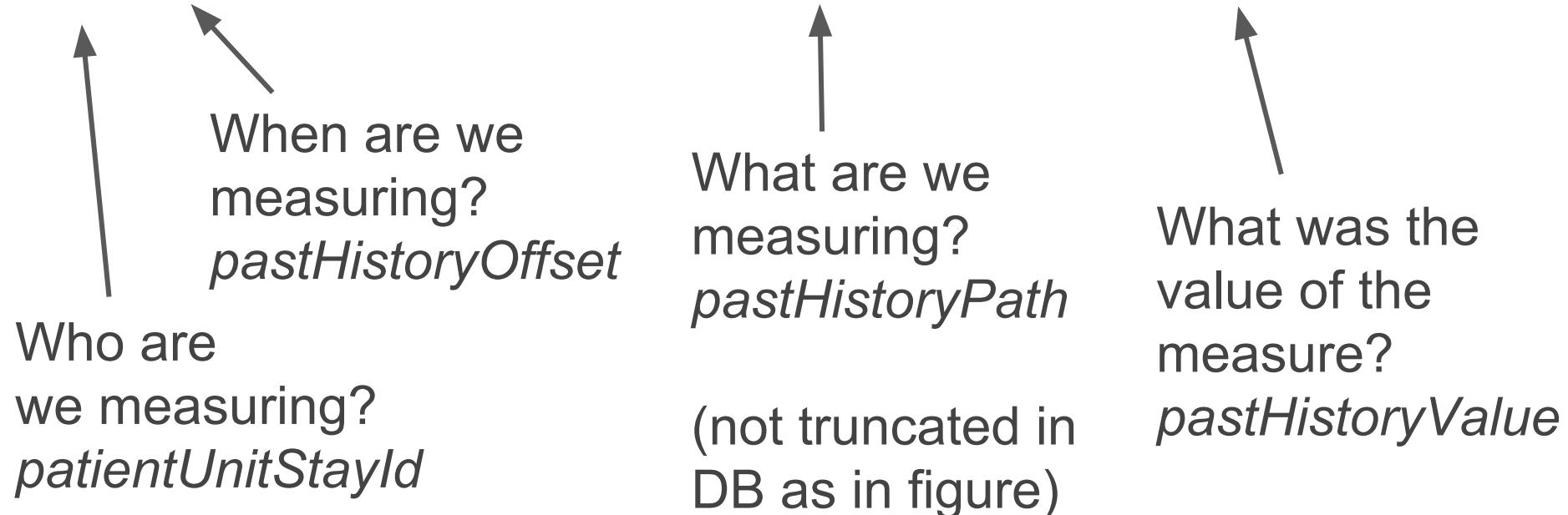
notes/Progress Notes/Past History/Past History Obtain Options/Performed
notes/Progress Notes/Past History/Organ Systems/Cardiovascular (R)/Congestive Heart Failure/CHF - class II
notes/Progress Notes/Past History/Organ Systems/Cardiovascular (R)/Valve disease/AS
notes/Progress Notes/Past History/Organ Systems/Cardiovascular (R)/Hypertension Requiring Treatment/hypertension requiring treatment
notes/Progress Notes/Past History/Organ Systems/Cardiovascular (R)/AICD/AICD



What are we
measuring?

Timestamped data - how most data is stored

pid	offset	time	time24	year	path	value
141168	60	evening	16:54:00	2015	notes/Progress Notes/Past History/Past History Obt	Performed
141168	60	evening	16:54:00	2015	notes/Progress Notes/Past History/Organ Systems/Ca	CHF - class II
141168	60	evening	16:54:00	2015	notes/Progress Notes/Past History/Organ Systems/Ca	AS
141168	60	evening	16:54:00	2015	notes/Progress Notes/Past History/Organ Systems/Ca	hypertension requiring treatment
141168	60	evening	16:54:00	2015	notes/Progress Notes/Past History/Organ Systems/Ca	AICD



Times - offsets instead of dates

- Most tables have an “offset” column - Minutes since ICU admission
 - Very useful! Defines the timing for all data.
- Other time related columns available in most tables...
 - Year
 - 2014, 2015
 - Time24
 - 10:00, 11:00, etc
 - Time
 - Midnight, noon, etc
 - *** this does not sensibly line up with Time24 (see next slide)

hour	midnight	night	evening	morning	midday	noon
00			834818			
01			805437			
02				770414		
03				724258		
04				698421		
05				735694		
06		867307				
07		1418370				
08		2370357				
09		3769600				
10						4632678
11						4593671
12						3648540
13						2527112
14	1705329					
15	1262547					
16	1125212					
17	1026952					
18					944782	
19					900860	
20					921850	
21					969608	
22			969336			
23			909378			

Inconsistent - not sure why!

Full path label

- Label is often stored with “path” or “string”
- These are hierarchical labels for the data
 - Often represent the data source + drop down menu
- **Diagnosis** - diagnosisstring
 - cardiovascular|ventricular disorders|congestive heart failure
- **pastHistory** - pasthistorypath
 - notes/Progress Notes/Past History/Organ Systems/Renal (R)/Renal Failure/renal failure- not currently dialyzed
- Sometimes pipe delimited, sometimes slash delimited



Care plans

Care plans - in order of data size

- carePlanEOL
- carePlanInfectiousDisease
- carePlanCareProvider
- carePlanGoal
- carePlanGeneral

carePlanEOL

- 0.6% of patients have an observation
- Times associated with a discussed end of life care
- No cell path
- Rarely populated

patientunitstayid	cpleolsaveoffset	cpleoldiscussionoffset	activeupondischarge
242505	5196	0	True
245984	1624	0	True
250634	92	-1079	True
251196	10049	8696	True
251924	4401	-921	True

carePlanInfectiousDisease

- 1.8% of patients have an observation
- Confirmed infections
- Communicate
 - site or source of infectious process
 - infection control precautions
- Very hospital specific
 - only some hospitals use it
- *infectDiseaseAssessment*
 - Possible infection
 - Probable Infection
 - Definite infection

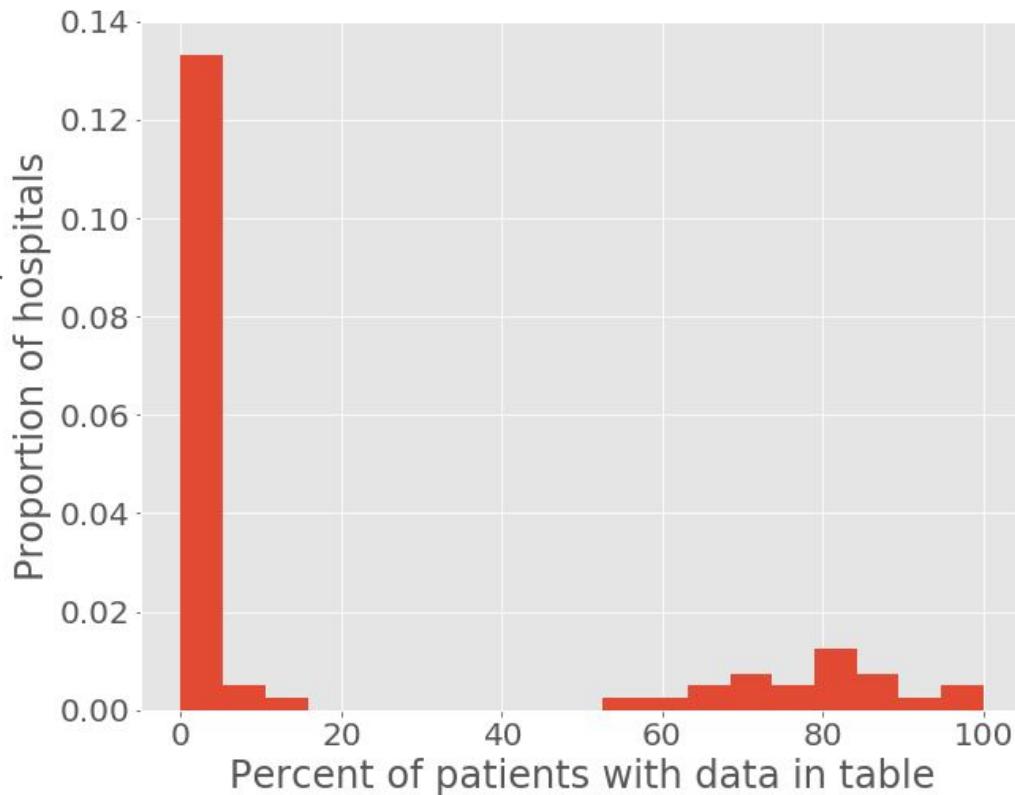
infectdiseasesite	count
Urinary tract	1791
CNS	91
Sinus	44
Disseminated	6
Skin & Soft tissue	450
Wound	216
Unknown	147
Bone	50
Prosthetic device	2
Catheter related bloodstream	12
Other	566
Lung	3078
Intra-abdominal	712
Blood	891

carePlanCareProvider

- 86% of patients have an observation
- *managingPhysician*
 - Consulting / Managing
- *specialty*
 - CCM (intensivist), cardiology, internal, hospitalist
- *activeUponDischarge*
 - True/false
- *providerType* - always null
- *interventionCategory* - I, II, III, IV, Unknown

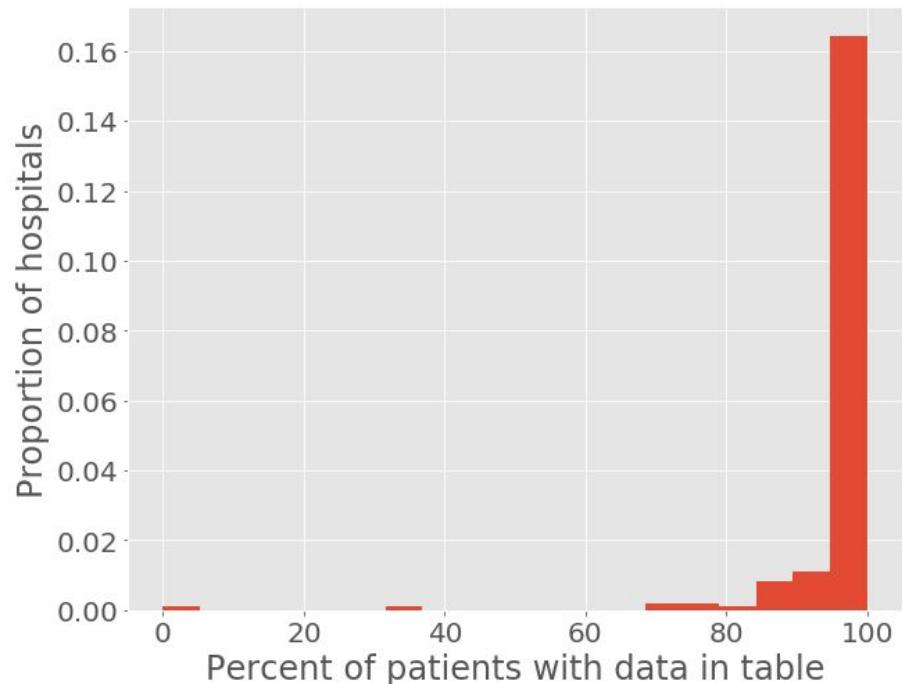
carePlanGoal

- 8.4% of patients have an observation, 30 obs per patient
- Goals
 - Hemodynamically stable within order parameters
 - Neurologically stable
 - Normal electrolytes
 - etc
- Absence of sepsis
- Absence of fever



carePlanGeneral

- 98% of patients have an observation
- ~15.78 obs per patient
- Structured input
 - Ventilation status
 - Code status
 - (e.g. stuff that should be in carePlanEOL)
 - Volume status (e.g. Hypovolemia)
 - Stress Ulcer Prophylaxis
 - Restraints
- Lots to discover!



Overview of tables

Patient info

*patient
hospital*

APACHE-IV

*admissionDx
apacheApsVar
apachePatientResult
apachePredVar*

Care plans

*carePlanCareProvider
carePlanEOL
carePlanGeneral
carePlanGoal
carePlanInfectiousDiseases*

Timestamped data

*admissionDrug
diagnosis
infusionDrug
lab
pastHistory
treatment
medication*

Monitor data

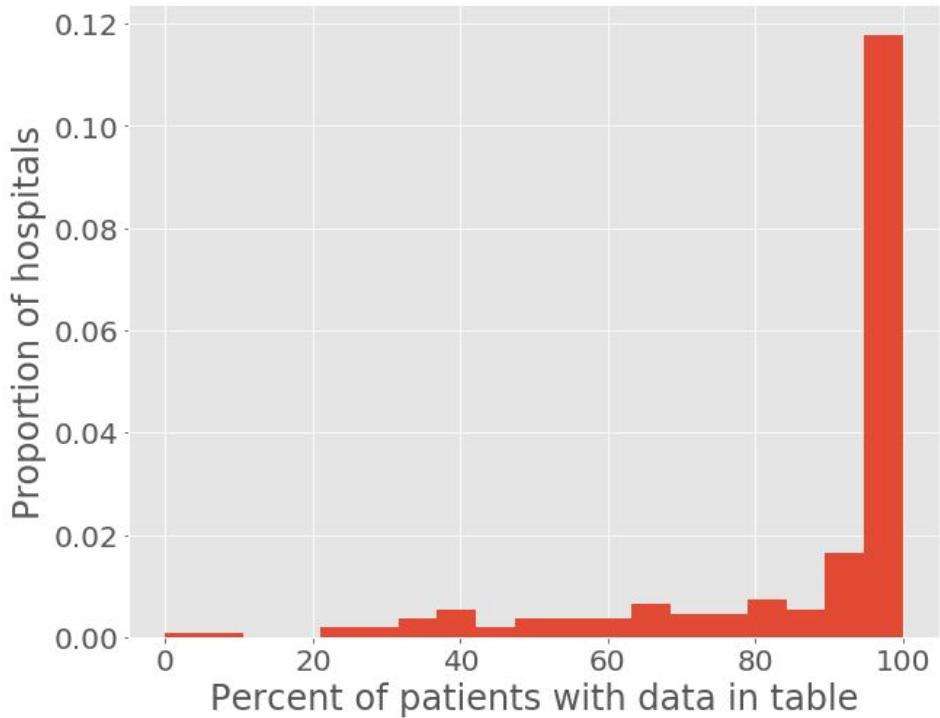
*vitalAperiodic
vitalPeriodic*



Timestamped data

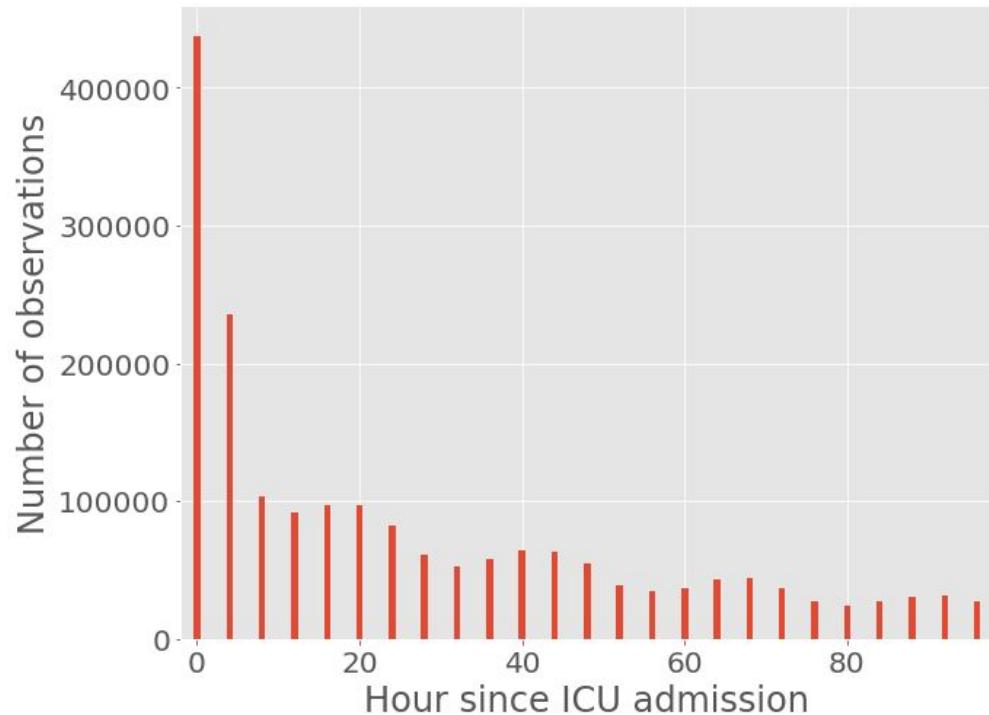
Diagnosis

- 86.2% have a diagnosis observation
- Actually an active problems list
 - Structured problem list - selected from a drop down menu
 - Organized by body system (Cardiovascular, Pulmonary)
- Primary, Secondary, Major



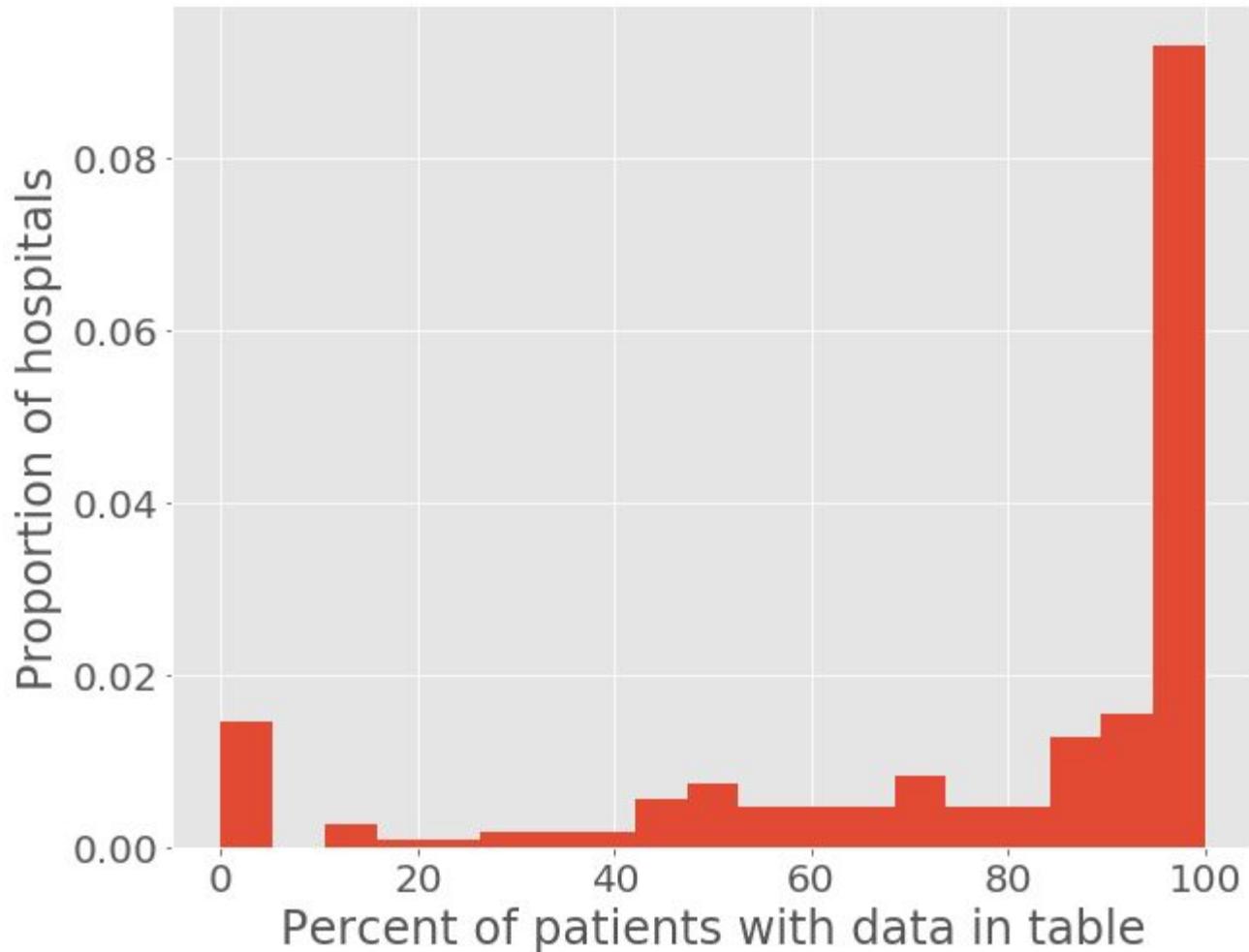
Diagnosis

- Usually present in the first 0-24 hours
- Less frequently populated later in the patient stay



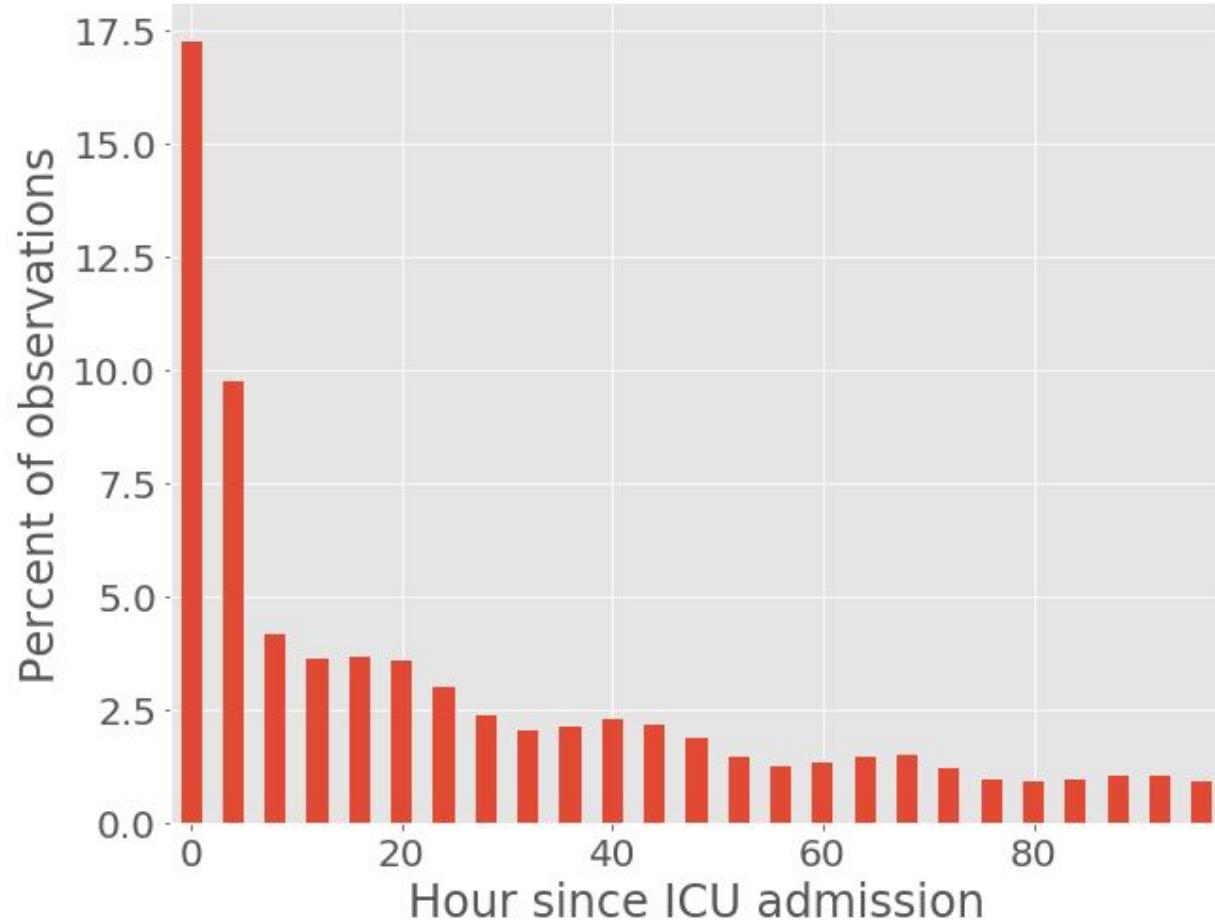
Treatment

- 75.9% of patients have a treatment observation
- Not used by every hospital
- Can identify use of vasopressors, ventilation, foley, insulin, etc



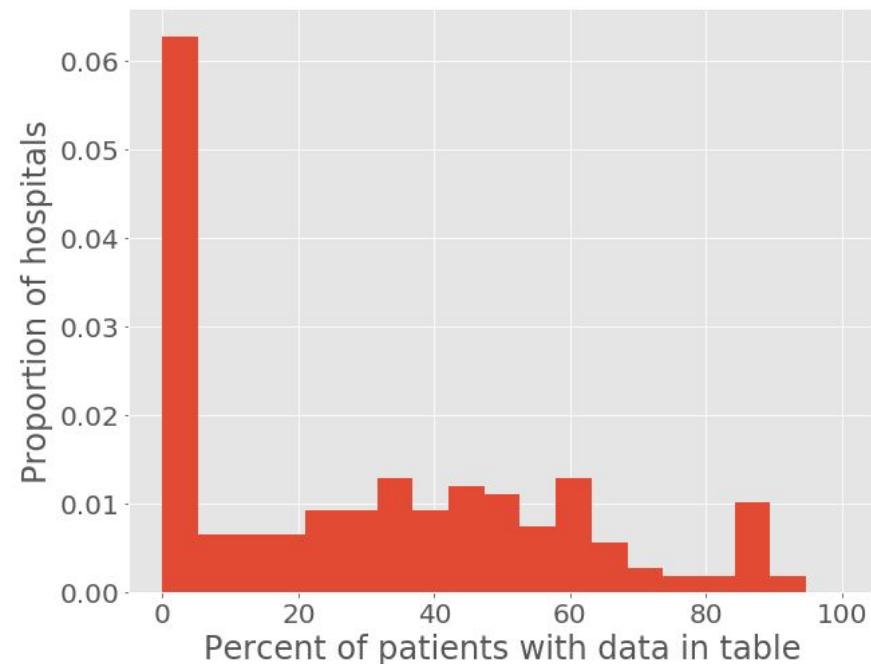
Treatment

- Most observations made during the first day



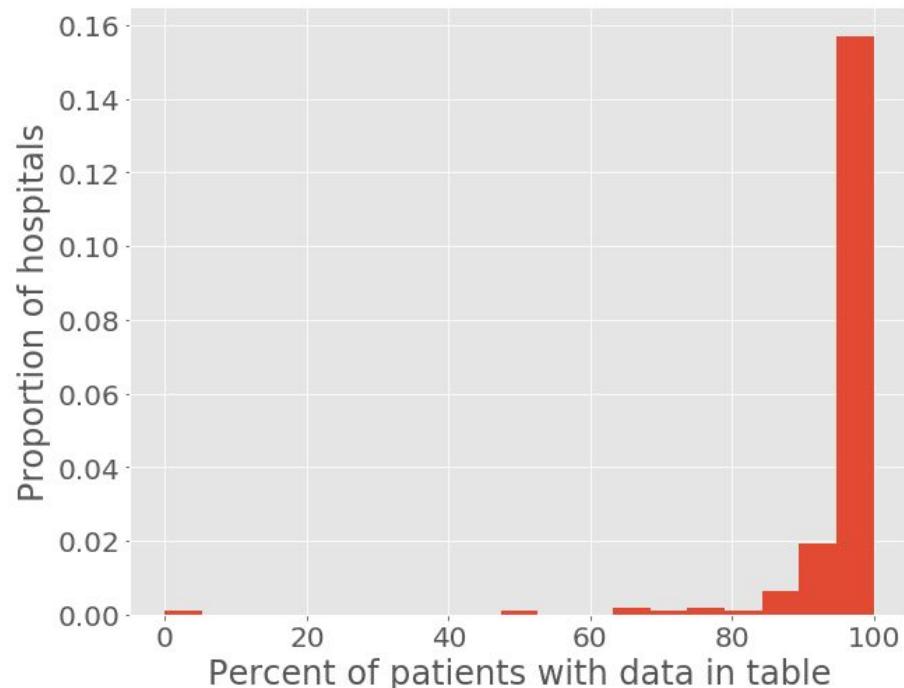
infusionDrug

- 36.6% of patients have an observation
- Need to develop methods on a hospital level to determine if data is present
- Some data charted manually
 - e.g. obs every 4-24 hours
- Some data sync from EHR
 - e.g. obs every 1 hour
- 3500+ unique drugs
 - E.g. “Dopamine” and “Dopamine ()”
 - Only ~200 are observed for more than 100 patients



Lab

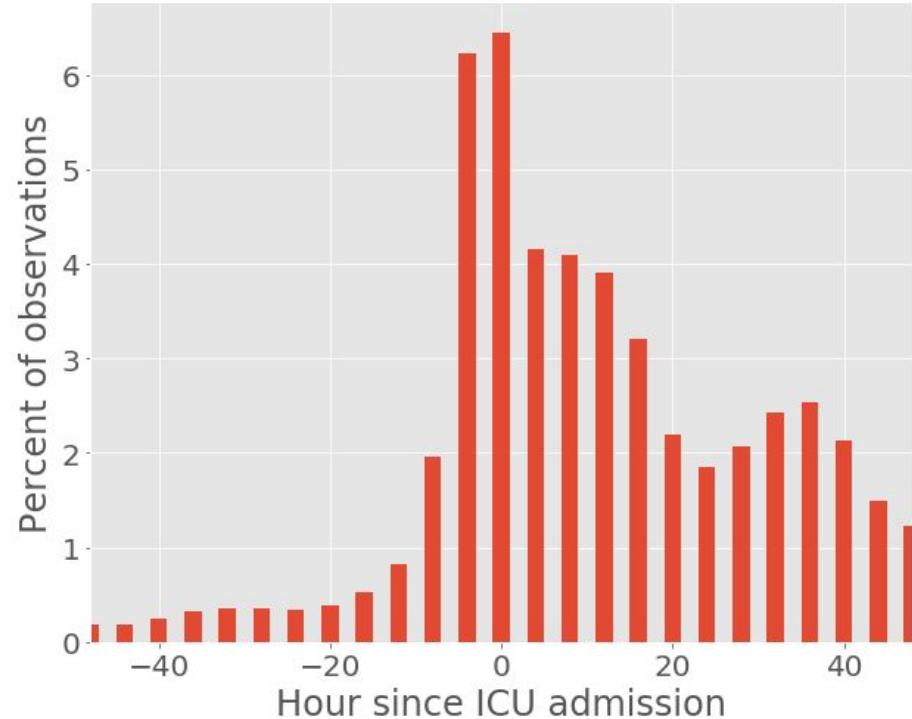
- Lab measurements from patient stays
- 97% of patients have an observation
- Labname - 158 unique labs
 - Structured!



patientunitstayid	labresultoffset	labname	labresult	labmeasurenamesystem
141168	2026	fibrinogen	177.0000	mg/dL
141168	1133	PT - INR	2.5000	ratio
141168	2026	magnesium	2.0000	mg/dL
141168	1133	PT	26.6000	sec
141168	2141	pH	7.2000	

Lab

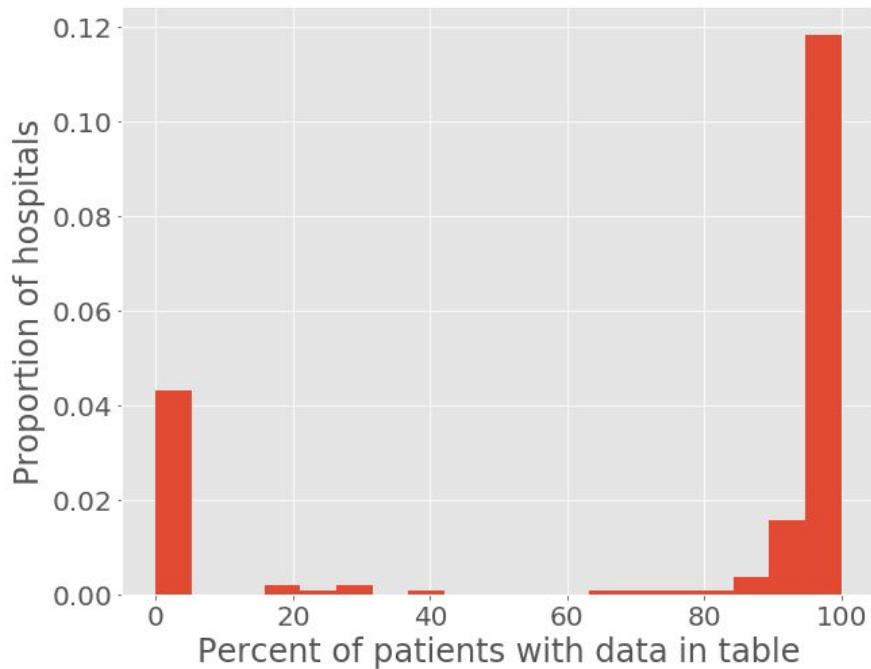
- Lab measurements from patient stays
- 97% of patients have an observation
- Labname - 158 unique labs
 - Structured!
- 20% have measurements more than 10 hours before ICU admission
 - Likely in-hospital labs



patientunitstayid	labresultoffset	labname	labresult	labmeasurenamesystem
141168	2026	fibrinogen	177.0000	mg/dL
141168	1133	PT - INR	2.5000	ratio
141168	2026	magnesium	2.0000	mg/dL
141168	1133	PT	26.6000	sec
141168	2141	pH	7.2000	

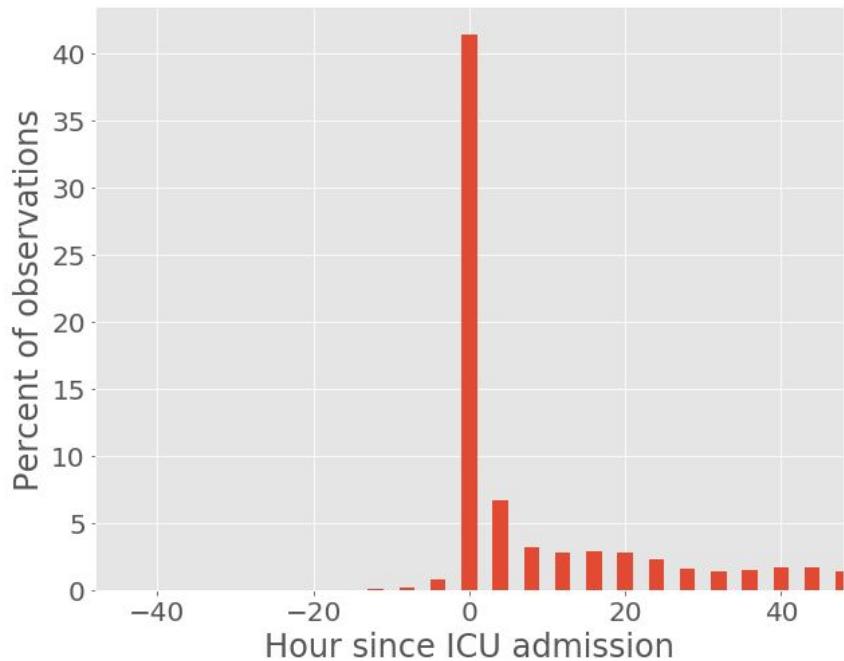
medication

- 82% of patients have an observation
- 44 observations per patient
- 71635 unique *drugName* :(
- Can use *drugHICLSeqNum* as ontology... but we don't have *d_hicl*

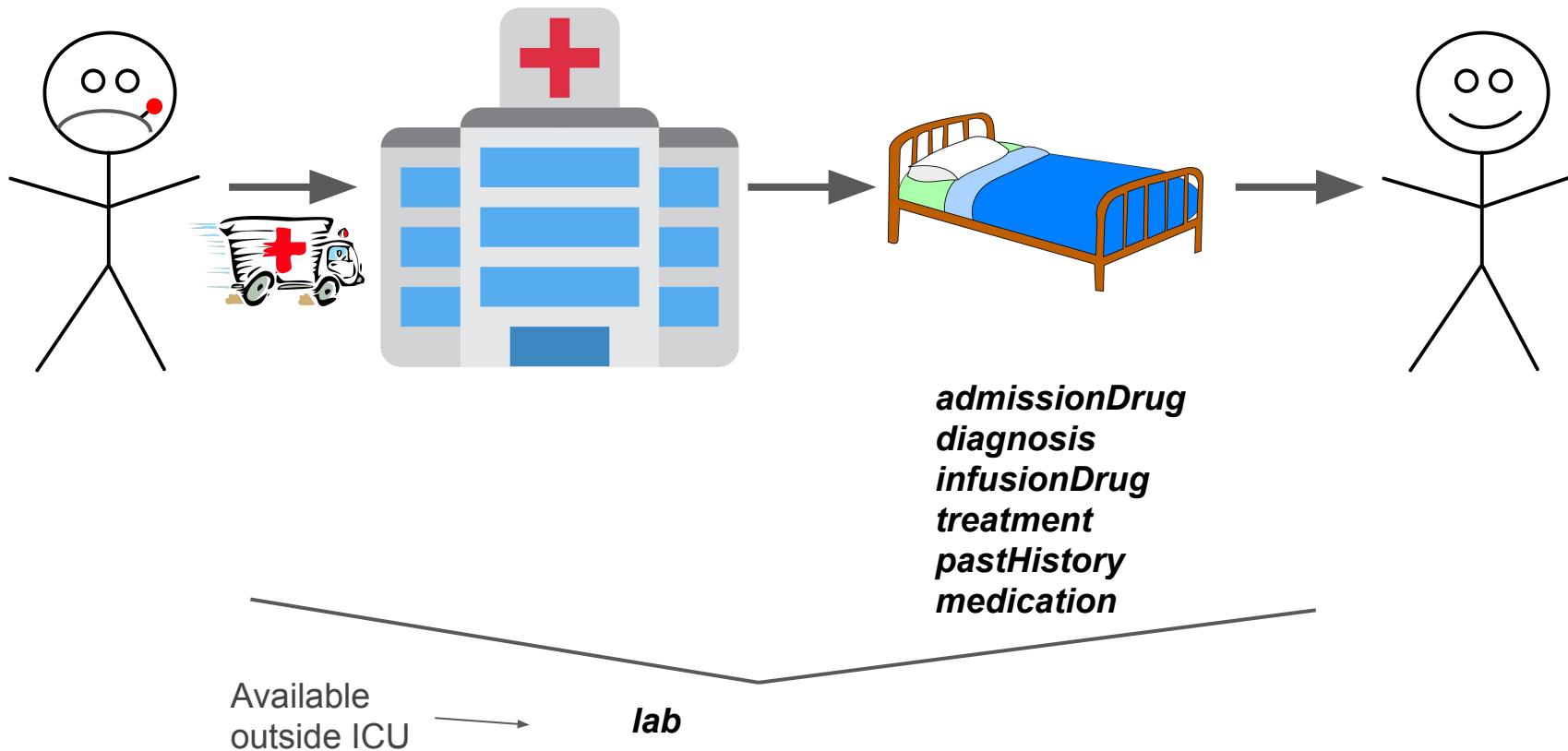


medication

- Probably only data from in ICU



Patient's hospital course



Overview of tables

Patient info

*patient
hospital*

APACHE-IV

*admissionDx
apacheApsVar
apachePatientResult
apachePredVar*

Care plans

*carePlanCareProvider
carePlanEOL
carePlanGeneral
carePlanGoal
carePlanInfectiousDiseases*

Timestamped data

*admissionDrug
diagnosis
infusionDrug
lab
pastHistory
treatment
medication*

Monitor data

*vitalAperiodic
vitalPeriodic*



Monitor data

vitalPeriodic

- 96% of patients have an observation
- Data from monitor - “unvalidated”
- One measurement every 5 minutes - big data!
- Common
 - Temperature, Heart rate, "sao2", Respiratory rate, ST levels
 - Invasive blood pressure - mean, systolic, diastolic
- Infrequent
 - Pulmonary artery pressure - mean, systolic, diastolic
 - CVP (central venous pressure)
 - End tidal CO₂, SVO₂, ICP (intracranial pressure)

vitalAperiodic

- 94.5% of patients have an observation
- *Aperiodic* - e.g. blood pressures every 4-6 hours
- Common
 - Non-invasive blood pressure - mean, systolic, diastolic
- Infrequent
 - Cardiac output, cardiac input
 - Pulmonary artery occlusion pressure (wedge pressure)
 - SVR, SVRI, PVR, PVRI

Overview of tables

Patient info

*patient
hospital*

APACHE-IV

*admissionDx
apacheApsVar
apachePatientResult
apachePredVar*

Care plans

*carePlanCareProvider
carePlanEOL
carePlanGeneral
carePlanGoal
carePlanInfectiousDiseases*

Timestamped data

*admissionDrug
diagnosis
infusionDrug
lab
pastHistory
treatment
medication*

Monitor data

*vitalAperiodic
vitalPeriodic*

Overview of data

MIMIC-III tables also listed



intakeOutput
nurseCharting
OUTPUTEVENTS

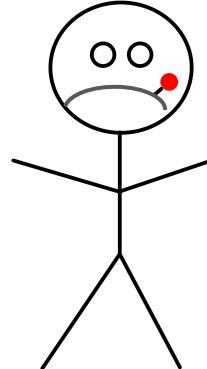
infusionDrug
intakeOutput
nurseCharting
INPUTEVENTS_CV
INPUTEVENTS_MV



carePlanGeneral, etc
<< NOTEVENTS



admissionDrug
medications
PRESCRIPTIONS



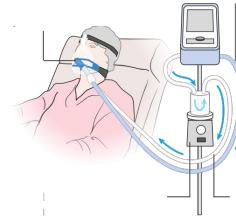
lab
LABEVENTS



apacheApsVar
apachePredVar
apachePatientResul
t
admissionDx



vitalPeriodic
vitalAperiodic



treatment
nurseCharting
CHARTEVENTS

Total row counts across tables

<i>admissiondrug</i>	874,920	<i>diagnosis</i>	2,710,672
<i>allergy</i>	251,949	<i>hospital</i>	459
<i>admissionDx</i>	626,858	<i>infusiondrug</i>	4,803,719
<i>apacheapsvar</i>	171,177	<i>lab</i>	39,132,531
<i>apachepatientresult</i>	297,064	<i>medication</i>	7,301,853
<i>apachepredvar</i>	171,177	<i>pasthistory</i>	1,149,180
<i>careplancareprovider</i>	502,765	<i>patient</i>	200,859
<i>careplaneol</i>	1,433	<i>treatment</i>	3,688,745
<i>careplangeneral</i>	3,115,018	<i>vitalperiodic</i>	25,075,074
<i>careplangoal</i>	504,139	<i>vitalperiodic</i>	146,671,642
<i>careplaninfectiousdisease</i>	8,056		

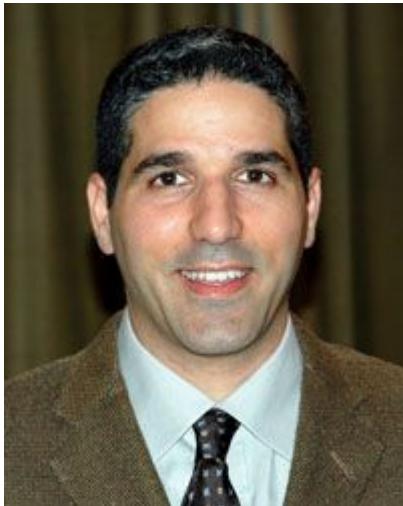


This is great and everything is perfect!

Except just this one thing.. maybe two.. well....

Issues about data will be posted publicly

<https://github.com/MIT-LCP/eicu-website/issues/34>



obadawi commented 5 days ago

Member +

We recently discovered that the ETL transforming raw eCM data into eRI data converts null date/time fields into "0" which will then appear to have an offset equal to the time of ICU admission. We believe this mainly affects the medications table which may often have null stop dates when medications are ordered but are trying to investigate further. This is something we can fix in a future eRI release but I doubt it can be updated in the current CRD release. We should add comments to the documentation to inform users of this and we can add more clarity to it as we learn more about what areas are affected.

Tables not released

- customLab
- intakeOutput
- microLab
- note (just quick comments)
- nurseAssessment
- nurseCare
- nurseCharting
- physicalExam
- respiratoryCare
- respiratoryCharting

The screenshot shows a dark-themed website for the eICU Collaborative Research Database. At the top, there's a navigation bar with links for 'About', 'Getting started', 'Data details', and 'Tables in eICU'. The 'Tables in eICU' link is highlighted with a red box. Below the navigation, there's a list of table names: 'admissionondrug', 'admissionondx', 'allergy', 'apacheApsVar', and 'apachePatientResult'. The main content area features a large heading 'nurseAssessment' and a sub-heading 'This table is not currently available in the public dataset.' Below this, there's a section titled 'Purpose:' followed by a detailed description of the table's function.

nurseAssessment

This table is not currently available in the public dataset.

Purpose:

The Nursing Assessment Flowsheet provides the capability to assess and document patient items such as pain, psychosocial status, patient/family education, neurologic, cardiovascular, respiratory, oral/GI/GU, skin, and other nursing assessment data along with date/time.

Unreleased tables are still documented online
<https://eicu-crd.mit.edu>

Number of rows in unreleased tables

Table Name	Rows	Patients	Patients (%)
customLab	1082	197	0.10
intakeOutput	12,994,887	181,002	90.11
microLab	16,996	2,923	1.46
note	3,835,897	178,303	88.77
nurseAssessment	18,255,000	13,003	6.47
nurseCare	13,618,762	13,054	6.50
nurseCharting	221,108,200	194,194	96.68
physicalExam	9,873,332	176379	87.81
respiratoryCare	865,381	44,772	22.29
respiratoryCharting	26,747,237	135,215	67.32

nurseCharting (unreleased) - example data

patientunitstayid	offset	category	label	value
141168	-594	Other Vital Signs and Infusions	Abnormality Type-Skin Abnormality Arm Left Abrasio	Abrasions/Scratches
141168	6	Other Vital Signs and Infusions	BP Location	RUE
141168	6	Other Vital Signs and Infusions	Heart Rate Source	Monitor
141168	6	Other Vital Signs and Infusions	MAP (mmHg)	67
141168	6	Other Vital Signs and Infusions	Morse Fall Scale Score	70
141168	6	Other Vital Signs and Infusions	Pulse Ox Mode	Continuous
141168	6	Vital Signs	Heart Rate	140
141168	6	Vital Signs	Non-Invasive BP	59
141168	6	Vital Signs	Non-Invasive BP	82
141168	21	Other Vital Signs and Infusions	Additional Parameters	Cough
141168	21	Other Vital Signs and Infusions	Alterations in Voiding	Incontinence-type unknown
141168	21	Other Vital Signs and Infusions	Apical Pulse	Tachycardic
141168	21	Other Vital Signs and Infusions	Breath Sounds - LLL	Diminished
141168	21	Other Vital Signs and Infusions	Breath Sounds-LUL	Inspiratory wheeze;Diminished;Expiratory wheeze
141168	21	Other Vital Signs and Infusions	Breath Sounds-RLL	Diminished
141168	21	Other Vital Signs and Infusions	Breath Sounds-RML	Diminished;Expiratory wheeze
141168	21	Other Vital Signs and Infusions	Breath Sounds-RUL	Expiratory wheeze;Diminished
141168	21	Other Vital Signs and Infusions	Central Perfusion	Cool
141168	561	Scores	Pain Score/Goal	0
141168	563	Other Vital Signs and Infusions	Heart Rate Source	Monitor
141168	563	Other Vital Signs and Infusions	Heart Rhythm	A Flu
141168	563	Other Vital Signs and Infusions	Pulse Ox Mode	Intermittent
141168	563	Vital Signs	Heart Rate	132
141168	563	Vital Signs	O2 Saturation	99
141168	563	Vital Signs	Respiratory Rate	24
141168	563	Vital Signs	Temperature	TEMPORAL ARTERY
141168	563	Vital Signs	Temperature	36.7
141168	563	Vital Signs	Temperature	98.1
141168	566	Other Vital Signs and Infusions	Devices and Equipment	IV pump - single;Pulse oximetry monitor
141168	566	Other Vital Signs and Infusions	Precautions	Fall precautions
141168	566	Other Vital Signs and Infusions	Skin Protection/Pressure Redistribution Devices	Heel off loading with pillow(s)
141168	566	Other Vital Signs and Infusions	Special Needs	IV Line;Incontinence;Oxygen;Pressure ulcer risk;
141168	566	Other Vital Signs and Infusions	Specialty Mattress/Bed	Bed-Integrated bed system



Planned improvements

Release more data

- No real timeline ... and a difficult task
- Likely order
 - intakeOutput
 - respiratoryCharting / respiratoryCare
 - nurseCharting
 - ... the rest

Partition the data appropriately

- PostgreSQL 10 has declarative partitioning
 - Saves a lot of effort in partitioning
- Recently repartitioned CHARTEVENTS (mimic-code postgres10 branch)
 - Query aggregating heart rate and respiratory rate for the first day
 - Original: 32 seconds
 - Partitioned: 5 seconds
- PostgreSQL 10 has parallel querying
 - Partitioned: 5 seconds
 - Parallel: 2 seconds
 - *** must use a where clause for planner to use parallel sequential scan

Create a “skinny” version

- Huge number of redundant columns
- Level 1 skinny - we may add this option to eicu-code build scripts
 - Drop superfluous columns
 - Shorten data types if it will make a difference (e.g. BIGINT -> INT)
 - No schema change :(
- Level 2 skinny - we would not do this, but it may be worthwhile
 - Normalize the database (e.g. STRING -> ITEMID)
 - Prevents collation of strings in sequential scan which is sloooooooooooooow
 - Would change the schema :(

Explore!

Schema: <https://mit-lcp.github.io/eicu-schema-spy/>

Tutorial: <https://github.com/MIT-LCP/ACM-BCB-2017>