Deep dive into clinical data

HST.956/6.793
Feb 14, 2023

https://www.perthnow.com.au/business/oil-gas/deep-sea-divers-injured-off-wa-coast-in-high-pressure-incident-ng-b88698620z

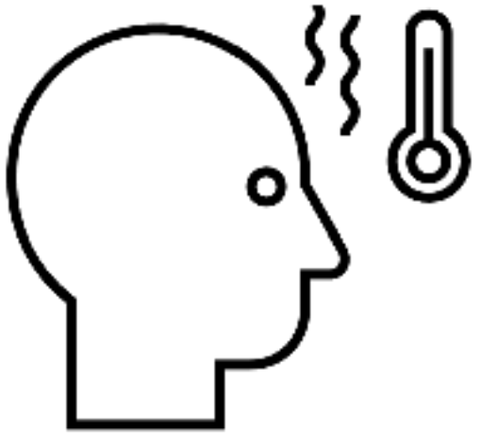
Massachusetts Institute of Technology

- Note: This slide deck contains a superset of the slides to be presented in class.

ML for Health Conferences

- Machine Learning for Health (ML4H) https://ml4health.github.io
  - Previously a NeurIPS workshop, separate symposium as of 2021
  - Last year, submissions due September, meeting in December
- Machine Learning for Healthcare (MLHC) https://www.mlforhc.org/
  - Submission deadline: TBD (probably April)
  - Columbia University, August 11-12, 2023
- Symposium on Artificial Intelligence for Learning Health Systems (SAIL) https://sail.health/
  - Submission deadline: Jan 20, 2023 (too late for 2023)
  - Puerto Rico, May 9–12, 2023
- Conference on Health, Inference, and Learning (CHIL) https://www.chilconference.org/
  - Submission deadline: Feb 15, 2023
  - Cambridge, MA, June 22-24, 2023
- And more (NeurIPS, ICML, AAAI, etc.)

Stakeholders in Healthcare



Patient

Providers

Payer

Policymaker

"The Four Ps" of healthcare

Stakeholders have different goals and expectations from the healthcare system

Policymaker

CENTERS FOR DISEASE CONTROL AND PREVENTION™

https://en.wikipedia.org/wiki/Centers_for_Disease_Control_and_Prevention
https://www.stetson.edu/other/safer-stetson/isolation.php

Overview of Clinical Data Science

- Topics of Discussion
  - Goals of Clinical Data Science
  - Sources of Clinical Data
  - Exploring Clinical Data
  - Challenges of Working with Clinical Data
  - Applying Clinical Data Science

Another Madhur Example: Mrs. Patel

65 year old female
Presents to the ER with abdominal pain
CT scan

https://radiopaedia.org/cases/renal-cell-carcinoma-9

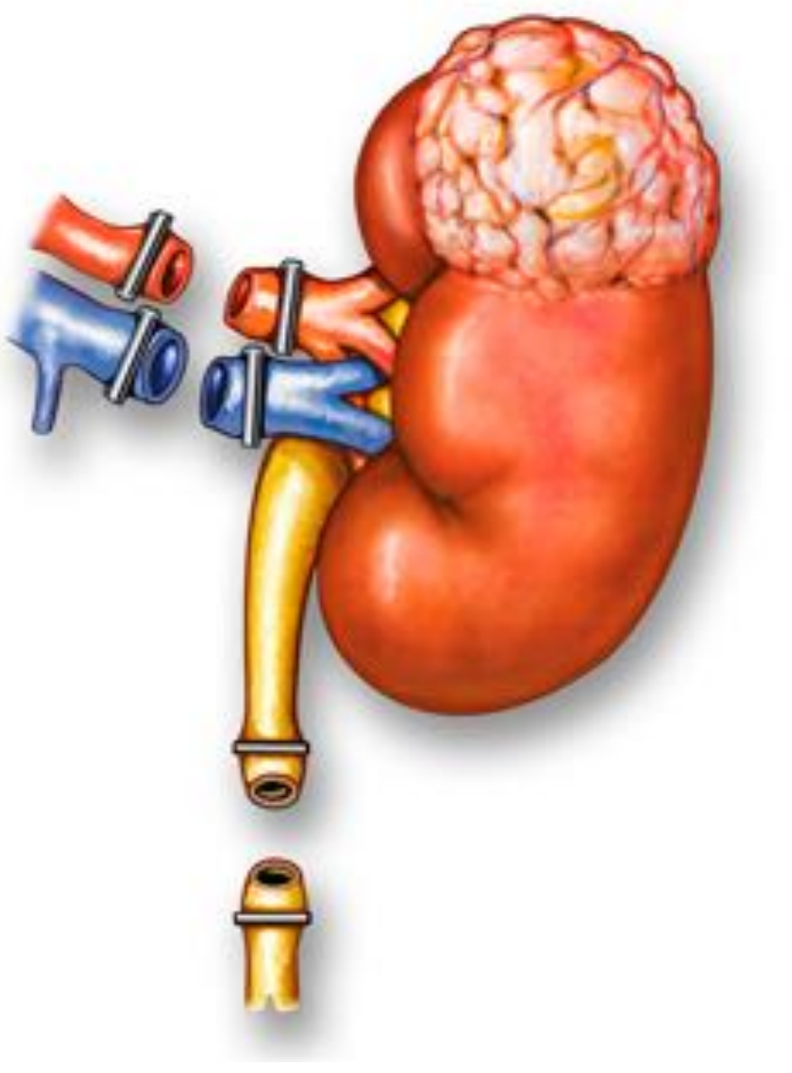
She is discharged from the ER and outpatient follow-up is arranged

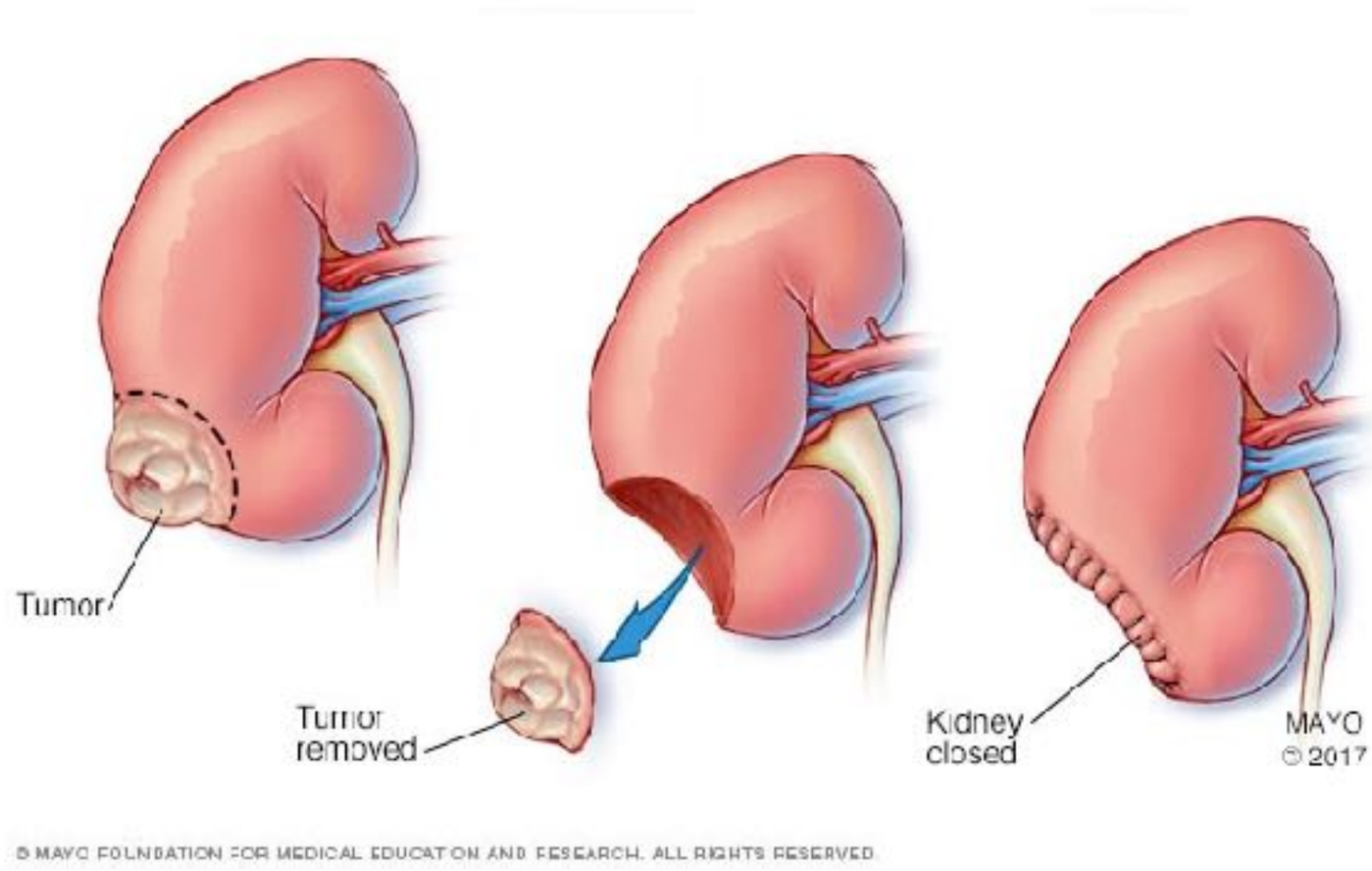
Case courtesy of Dr Roberto Schubert, Radiopaedia.org, rID: 14439

Patient/Provider Goals of Clinical Data Science

- Mrs. Patel is a 65 year old who was recently diagnosed with kidney cancer. She presents to your office. You discuss the diagnosis and treatment options. She has some questions.
  - After treatment, **what is the risk** of my cancer coming back before the Ultimate World Cruise (December 2023)?
  - **Will the risk** of my cancer coming back **change** if I get a partial nephrectomy instead of a radical nephrectomy?

Radical Nephrectomy VS Partial Nephrectomy

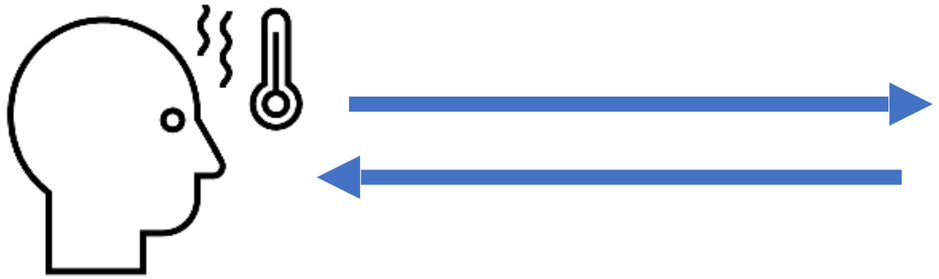




© MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. ALL RIGHTS RESERVED.

https://www.fairbanksurology.com/robotic-radical-nephrectomy
https://www.mayoclinic.org/tests-procedures/nephrectomy/multimedia/img-20332175

Sources of Clinical Data



Patient

Providers

Healthcare professionals

Hospital

Payer

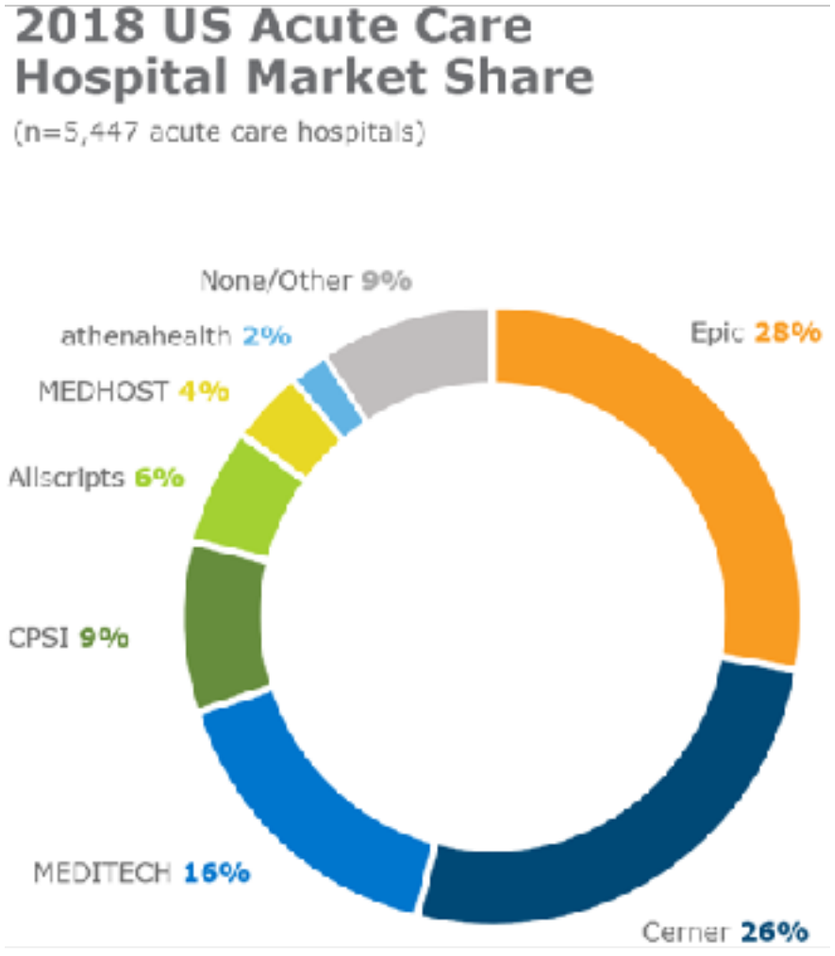
Policymaker

Provider Derived Data

- Previously, paper charts were used for clinical documentation.
  - What are problems with paper charts?
- Electronic health records (EHR) are a digital version of the patient's paper chart.
  - Providers are reimbursed based on the EHR
- Examples of EHR databases: MIMIC, Mass General Brigham Research Patient Data Registry (RPDR)

Electronic Health Records in the US

- Different hospitals use different EHR systems
  - Largest EHR systems in US
    1. EPIC
    2. Cerner
    3. Meditech
- To efficiently and accurately share clinical information, EHRs must be interoperable
  - Current EHRs are not interoperable

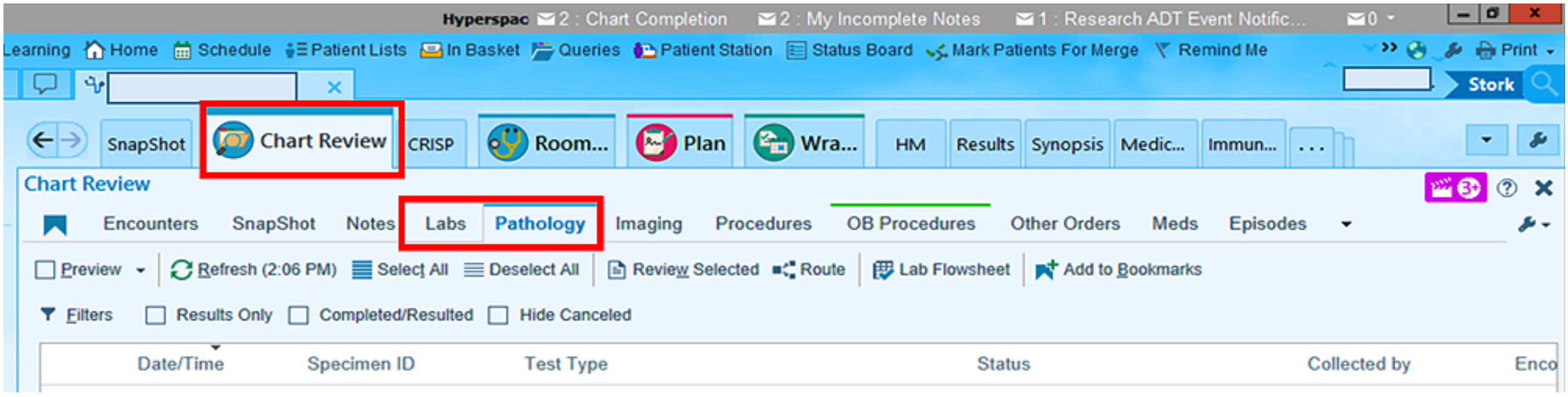


EPIC

How to find Pathology Results and Reports on EPIC

During MS3-MS4 clerkships, try to read the patients' labs and pathology notes within JHH on Epic:

Access **Chart Review ->** click "**Labs**" for microbiology testing or "**Pathology**" for diagnostic results

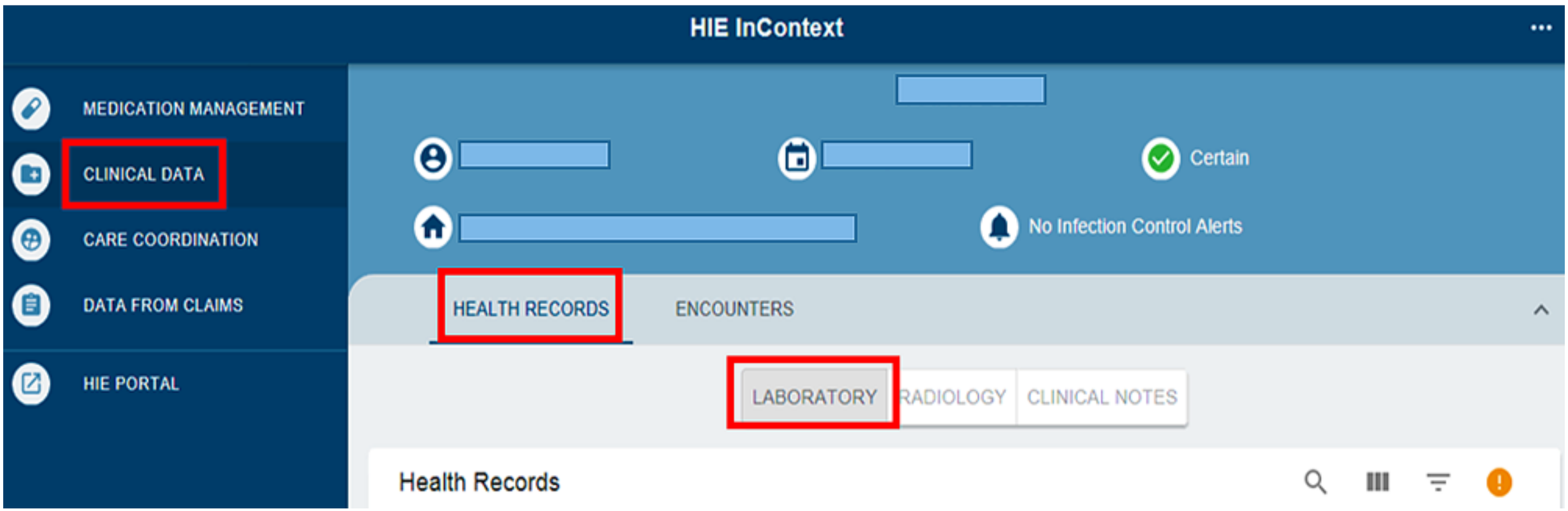


http://apps.pathology.jhu.edu/team-path-md/pathology-for-core-clinical-clerkships/how-to-find-pathology-results-and-reports-on-epic/

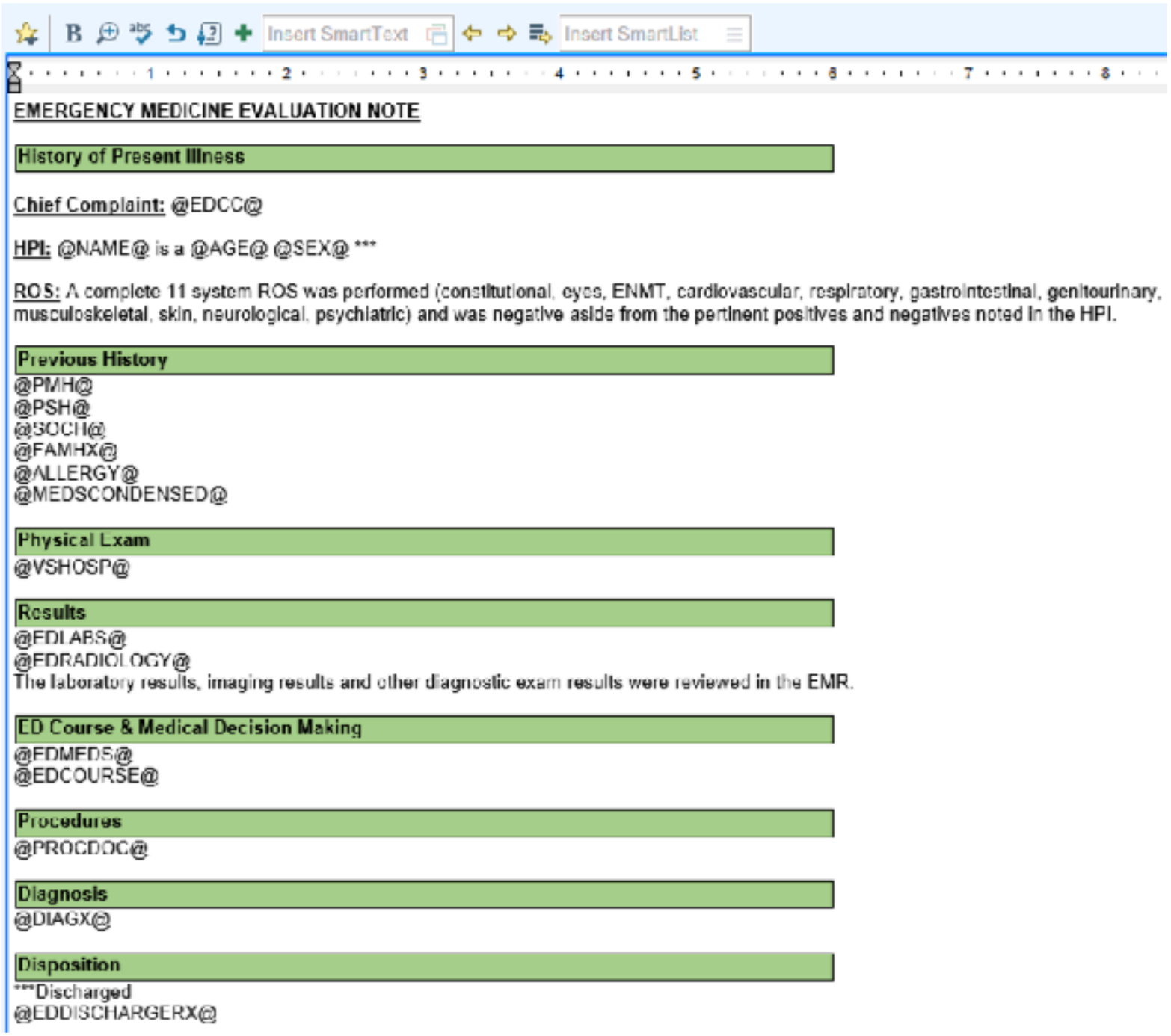
EPIC

If the lab tests were performed outside of JHH's hospital system:

Access **CRISP ->** click **Clinical Data -> Health Records -> Laboratory** for molecular tests



http://apps.pathology.jhu.edu/team-path-md/pathology-for-core-clinical-clerkships/how-to-find-pathology-results-and-reports-on-epic/

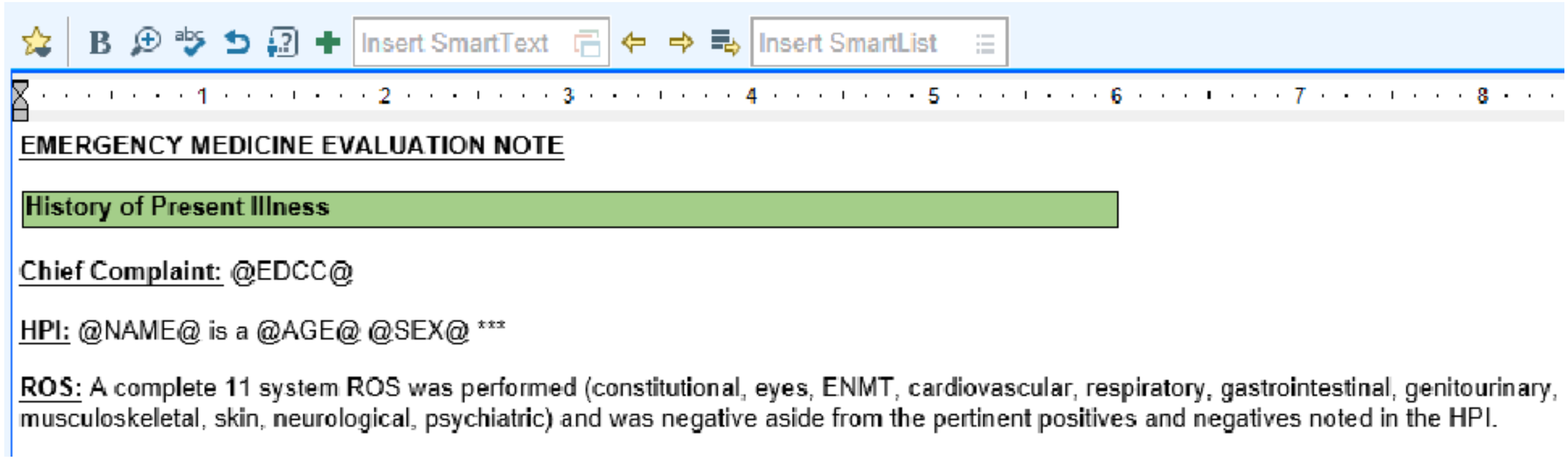


Sections of templated note

- History of Present Illness
- Previous History
- Physical Exam
- Results
- ED Course & Medical Decision Making
- Procedures
- Diagnosis
- Disposition

https://www.acep.org/administration/quality/health-information-technology/epic-articles/things-you-can-do-on-your-own-epic/

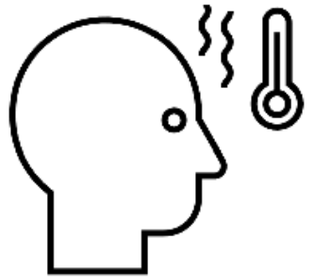
Trouble with Templates



- Is the Review Of Systems correct?
  - Would you perform a GI ROS for someone with a sprained ankle?
- Pre-populations of fields with old data
  - E.g., medication list is often out of date

https://www.acep.org/administration/quality/health-information-technology/epic-articles/things-you-can-do-on-your-own-epic/

Sources of Clinical Data

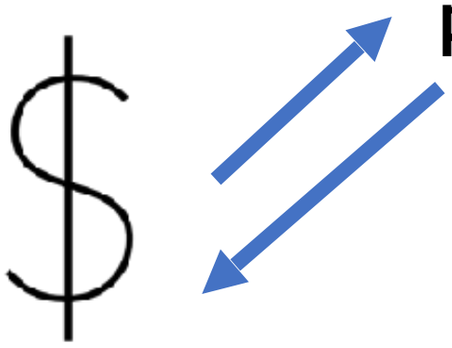


Patient

Providers

Hospital

Healthcare professionals

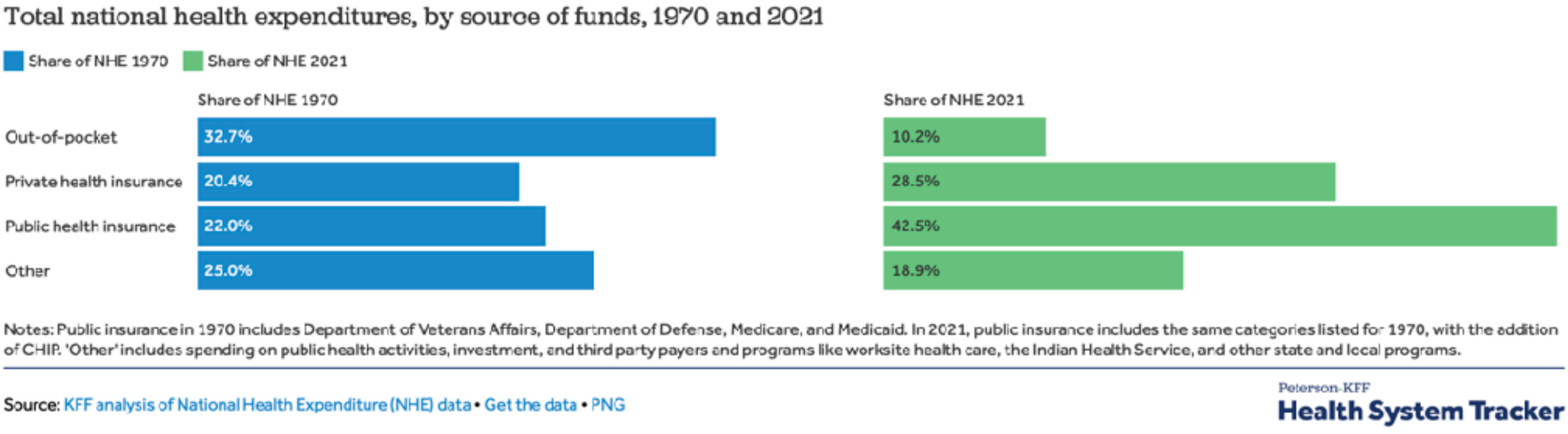


Payer

Policymaker

Payer Derived Data

- Claims data
  - Consists of the **billing codes** that **providers** (physicians, hospitals, pharmacies, and other health care providers) submit to **payers**
  - Examples: IQVIA, IBM Marketscan, Optum, Medicare
- Medicare Claims Data
  - Medicare is a Federal health insurance program
  - Covers: Age ≥ 65; Certain people under 65 with disabilities; People of any age with End Stage Renal Disease or amyotrophic lateral sclerosis



Payer Derived Data
CMS1500

HEALTH INSURANCE CLAIM FORM

APPROVED BY NATIONAL UNIFORM CLAIM COMMITTEE (NUCC) 02/12

ABC Insurance Company
Suite 600
567 Insurance Lane
Big City IL 80605

PICA

1. MEDICARE (Medicare#)	MEDICAID (Medicaid#)	TRICARE (ID#/DoD#)	CHAMPVA (Member ID#)	GROUP HEALTH PLAN (ID#)	FECA BLK LUNG (ID#)	OTHER (ID#)
				X		

1a. INSURED'S I.D. NUMBER (For Program in Item 1): X0123456789

2. PATIENT'S NAME (Last Name, First Name, Middle Initial)	3. PATIENT'S BIRTH DATE MM DD YY	SEX	4. INSURED'S NAME (Last Name, First Name, Middle Initial)
Doe Jr. John, J	01 01 1987	M X F	Doe. John, J

5. PATIENT'S ADDRESS (No., Street)	6. PATIENT RELATIONSHIP TO INSURED	7. INSURED'S ADDRESS (No., Street)
123 Main Street	Self Spouse Child X Other	123 Main Street

CITY	STATE	8. RESERVED FOR NUCC USE	CITY	STATE
Anytown	IL		Anytown	IL

ZIP CODE	TELEPHONE (Include Area Code)	ZIP CODE	TELEPHONE (Include Area Code)
60610	(312) 5551212	60610	(312) 5551212

9. OTHER INSURED'S NAME (Last Name, First Name, Middle Initial)	10. IS PATIENT'S CONDITION RELATED TO:	11. INSURED'S POLICY GROUP OR FECA NUMBER
Doe. Mary. A		A1234
a. OTHER INSURED'S POLICY OR GROUP NUMBER: X9876543210	a. EMPLOYMENT? (Current or Previous) YES X NO	a. INSURED'S DATE OF BIRTH MM DD YY: 01 01 1958 SEX M X F
b. RESERVED FOR NUCC USE	b. AUTO ACCIDENT? YES X NO PLACE (State)	b. OTHER CLAIM ID (Designated by NUCC): Y4 112233445566
c. RESERVED FOR NUCC USE	c. OTHER ACCIDENT? YES X NO	c. INSURANCE PLAN NAME OR PROGRAM NAME: ABC Insurance Company
d. INSURANCE PLAN NAME OR PROGRAM NAME: XYZ Insurance Company	10d. CLAIM CODES (Designated by NUCC)	d. IS THERE ANOTHER HEALTH BENEFIT PLAN? X YES NO If yes, complete items 9, 9a, and 9d.

READ BACK OF FORM BEFORE COMPLETING & SIGNING THIS FORM.

12. PATIENT'S OR AUTHORIZED PERSON'S SIGNATURE I authorize the release of any medical or other information necessary to process this claim. I also request payment of government benefits either to myself or to the party who accepts assignment below.
SIGNED: Signature on File DATE: 09/30/12

13. INSURED'S OR AUTHORIZED PERSON'S SIGNATURE I authorize payment of medical benefits to the undersigned physician or supplier for services described below.
SIGNED: SOF

CARRIER

PATIENT AND INSURED INFORMATION

https://fiachraforms.com/shop/1500-02-12-standard-paper-claim-form/

Payer Derived Data
CMS1500

Diagnosis codes

Procedu

14. DATE OF CURRENT ILLNESS, INJURY, or PREGNANCY (LMP) MM DD YY: 09 30 2012 QUAL 431
15. OTHER DATE QUAL 454 MM DD YY: 09 25 2012
16. DATES PATIENT UNABLE TO WORK IN CURRENT OCCUPATION MM DD YY: FROM 09 25 2012 TO 10 28 2012
17. NAME OF REFERRING PROVIDER OR OTHER SOURCE: DN Jane A Smith MD
17a. G2 ABC1234567890
17b. NPI 0123456789
18. HOSPITALIZATION DATES RELATED TO CURRENT SERVICES MM DD YY: FROM 09 25 2012 TO 09 28 2012
19. ADDITIONAL CLAIM INFORMATION (Designated by NUCC)
20. OUTSIDE LAB? [X] YES [ ] NO $ CHARGES: 112500 00
21. DIAGNOSIS OR NATURE OF ILLNESS OR INJURY Relate A-L to service line below (24E) ICD Ind. 9
A. 998.59 B. 780.6 C. V18.0 D. E878.8
E. F. G. H. I. J. K. L.
22. RESUBMISSION CODE: 7 ORIGINAL REF. NO.: ABC12334567890
23. PRIOR AUTHORIZATION NUMBER

24. A. DATE(S) OF SERVICE From MM DD YY	To MM DD YY	B. PLACE OF SERVICE	C. EMG	D. PROCEDURES, SERVICES, OR SUPPLIES (Explain Unusual Circumstances) CPT/HCPCS	MODIFIER	E. DIAGNOSIS POINTER	F. $ CHARGES	G. DAYS OR UNITS	H. EPSDT Family Plan	I. ID. QUAL	J. RENDERING PROVIDER ID. #
1 09 30 12	09 30 12	11	Y	99241	25	ABCD	50 00	2	Y	G2	Z5678901234
										NPI	9876543210
2 10 01 11	01 01 11	11	N	A6410	P2	ABSS	45 00	2	N		12345678901
										NPI	0123456789
3										NPI	
4										NPI	
5										NPI	
6										NPI	

PHYSICIAN OR SUPPLIER INFORMATION

25. FEDERAL TAX I.D. NUMBER SSN [ ] EIN [X]
26. PATIENT'S ACCOUNT NO.: 12341234
27. ACCEPT ASSIGNMENT? (For govt. claims, see back) [X] YES [ ] NO
28. TOTAL CHARGE $ 190 00
29. AMOUNT PAID $
30. Rsvd for NUCC Use
31. SIGNATURE OF PHYSICIAN OR SUPPLIER INCLUDING DEGREES OR CREDENTIALS (I certify that the statements on the reverse apply to this bill and are made a part thereof.)
Joe Smith MD
09/30/12
SIGNED DATE
32. SERVICE FACILITY LOCATION INFORMATION
General Hospital
9876 Hospital Street
Anytown IL 60610-9876
a. 567891234 b. G2A1234567890
33. BILLING PROVIDER INFO & PH # (312) 5552222
Physician Practice Inc
1234 Healthcare Street
Anytown IL 60610-1234
a. 9876543210 b. G2Z5678901234

NUCC Instruction Manual available at: www.nucc.org PLEASE PRINT OR TYPE APPROVED OMB-0938-1197 FORM 1500 (02-12)

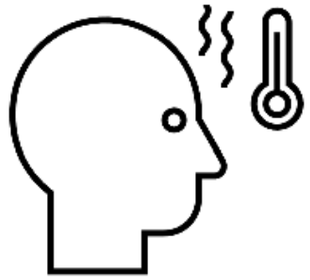
https://fiachraforms.com/shop/1500-02-12-standard-paper-claim-form/

Payer Derived Data

- All-payer claims databases
  - Large State databases that include claims from private and public payers
  - Massachusetts All-Payer Claims Database
    - Releases data extracts to government agencies, payers, providers, provider organizations, and researchers
    - All applications to access the data are reviewed for conformity with legal requirements
- What's missing?
  - E.g., test or procedure results
  - Outcomes

https://www.ahrq.gov/data/apcd/index.html
https://www.chiamass.gov/ma-apcd/

Sources of Clinical Data



Patient

Providers

Hospital

Healthcare professionals

Payer

Policymaker

Policymaker Derived Data

- **National Cancer Database (NCDB)**
  - Hospital registry data from **Commission on Cancer (CoC)**-accredited facilities
    - What is COC?
      - A program, from the American College of Surgeons, that **recognizes cancer care programs** for providing comprehensive, high-quality, and multidisciplinary patient centered care.
    - CoC accreditation
      - Granted to facilities that demonstrate compliance

CoC Standards

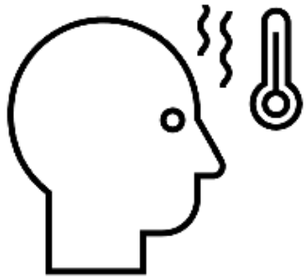
https://www.facs.org/-/media/files/quality-programs/cancer/coc/optimal_resources_for_cancer_care_2020_standards.ashx

CoC Standards

- **5.1 College of American Pathologists Synoptic Reporting**
- **Definition and Requirements: 90% of the eligible cancer pathology reports are structured using synoptic reporting** format as defined by the College of American Pathologists (CAP) cancer protocols, including containing all core data elements within the synoptic format.

https://www.facs.org/-/media/files/quality-programs/cancer/coc/optimal_resources_for_cancer_care_2020_standards.ashx

Sources of Clinical Data



Patient

Providers

Hospital

Healthcare professionals

Payer

Policymaker

Patient Derived Data

- PatientsLikeMe
  - Online community that allows members to find other patients like them, share and track their health data over time, and contribute to scientific research
  - Launched in 2006 for patients with amyotrophic lateral sclerosis
  - For-profit company
  - More than 600,000 registered members across more than 2900 conditions (as of February 2018)
    - Survey of members in 2016-2017
      - 67% furthered their understanding of how their condition could affect them
      - 63% on how to live better with their condition

Wicks, Paul, et al. "Scaling PatientsLikeMe via a “generalized platform” for members with chronic illness: web-based survey study of benefits arising." *Journal of medical Internet research* 20.5 (2018): e9909.

Patient Derived Data

Renal cell carcinoma

patientslikeme

HOME | CONDITIONS | TREATMENTS | SYMPTOMS | SEARCH | Sign in | Join now

Members are tracking more than 2,800 conditions on PatientsLikeMe. See what they're saying about yours...

Cancer
Breast , Lung , Liver , Testicular , Prostate , Pancreatic , CLL (Chronic Lymphocytic Leukemia) , Non-Hodgkin's Lymphoma , Thyroid

Developmental and Chromosomal
Tay-Sachs , Autism Spectrum , Down Syndrome

Digestive and Intestinal
Crohn's Disease , IBS , Ulcerative Colitis

Endocrine
Diabetes: Type I , Type II , Hypothyroidism , Hyperthyroidism

Eye, Ear, Nose and Throat
Hearing Loss , Glaucoma , Macular Degeneration

Heart, Blood and Circulatory
Coronary Artery Disease , Hypertension , Iron Deficiency Anemia , Raynaud's Syndrome , Congestive Heart Failure , Cardiomyopathy , Aplastic Anemia

https://www.patientslikeme.com/conditions/

Patient Derived Data

Common symptoms reported by people with renal cell cancer

Common symptoms	How bad it is	What people are taking for it
Pain		Pregabalin, Gabapentin, Oxycodone
Fatigue		phetamine, Armodafinil, Motorized scooter/chair
Stress		Aromatherapy
Anxious mood		Clonazepam, Escitalopram, Acupuncture
Depressed mood		Venlafaxine, Sertraline, Aripiprazole

21 renal cell cancer patients report severe pain (17%)

Reports may be affected by other conditions and/or medication side effects. We ask about general symptoms (anxious mood, depressed mood, fatigue, pain, and stress) regardless of condition.

Last updated: February 7, 2022

https://www.patientslikeme.com/conditions/renal-cell-ca

What is Venlafaxine?

- An antidepressant in a group of drugs called selective serotonin and norepinephrine reuptake inhibitors (SSNRIs).
- Affects chemicals in the brain that may become unbalanced and cause depression.
- Used to treat major depressive disorder, anxiety, and panic disorder.

https://www.patientslikeme.com/treatment/venlafaxine

Patient Derived Data

117 patient evaluations for Venlafaxine

Sep 3, 2012 (Started Oct 10, 2006)

Effectiveness	■■■□	Moderate (for major depressive disorder)
Effectiveness	■■■□	Moderate (for depressed mood)
Side effects	■■□□	Mild (for Overall) (sexual dysfunction)
Adherence	■■■■	Always
Burden	■□□□	Not at all hard to take

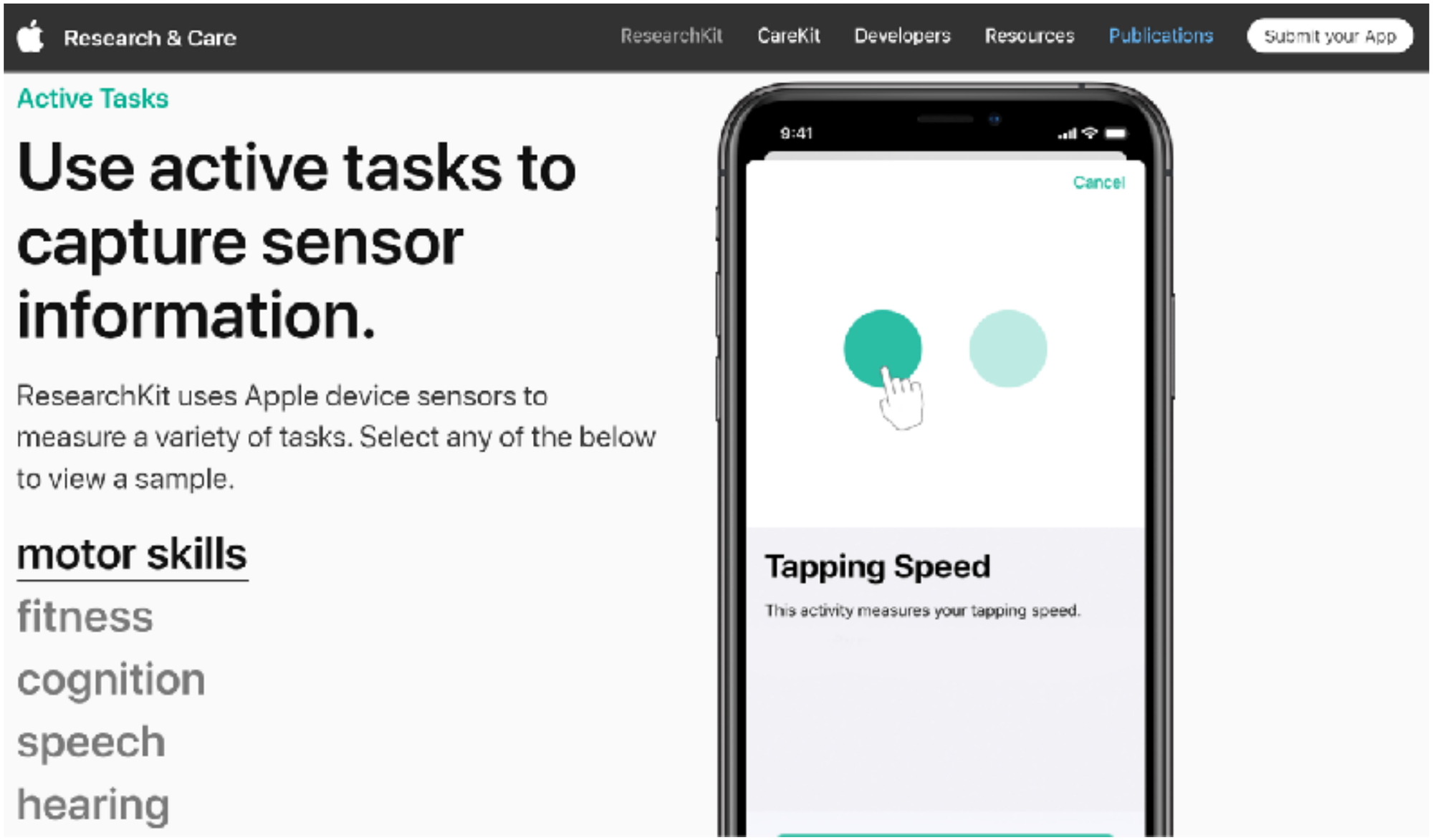
Dosage: 100 mg Daily

Advice & Tips: Slight sexual dysfunction. As long as I take it several hours before sexual activity it is no problem. A big benefit is the leveling out of emotions.

Cost: < $25 monthly

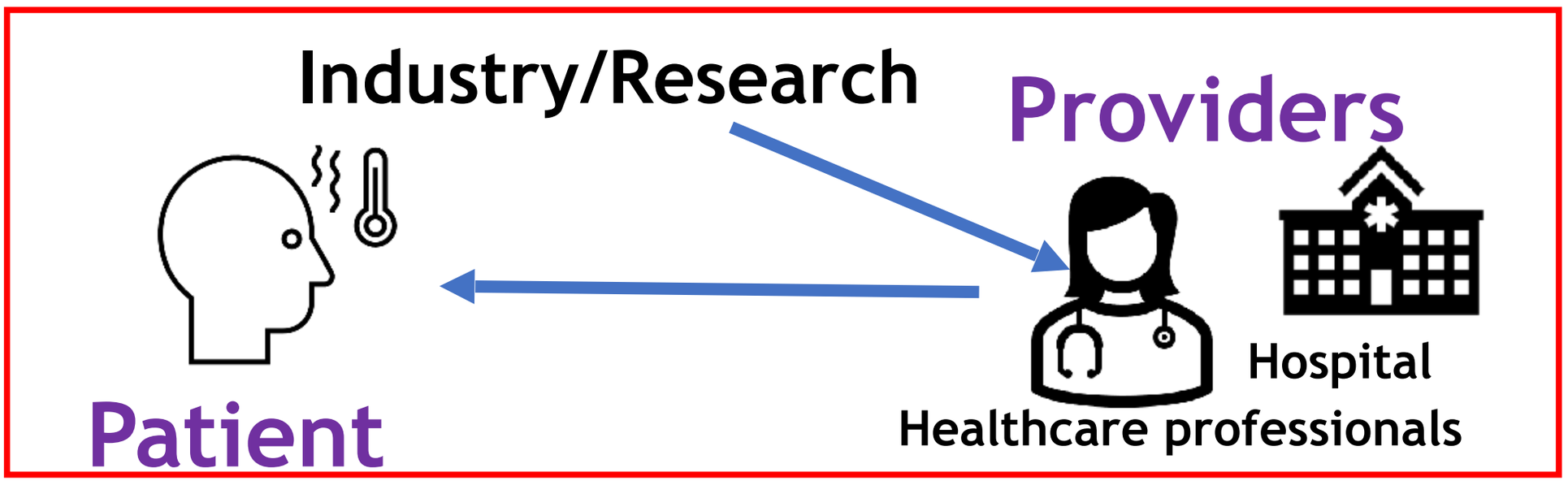
https://www.patientslikeme.com/treatment/duloxetine

Patient Derived Data



https://www.researchandcare.org/researchkit/

Sources of Clinical Data



- Study evaluating a new test or treatment
- May also be used for secondary analyses

Payer

Policymaker

Where does clinical data come from?

- Patient
- Providers
- Payer
- Policy-maker
- Industry
- Research

Caution

- None has complete data for
  - Individual patient
  - Population
- Usually not designed for research

Patient/Provider Goals of Clinical Data Science
(now that we know what data we have)

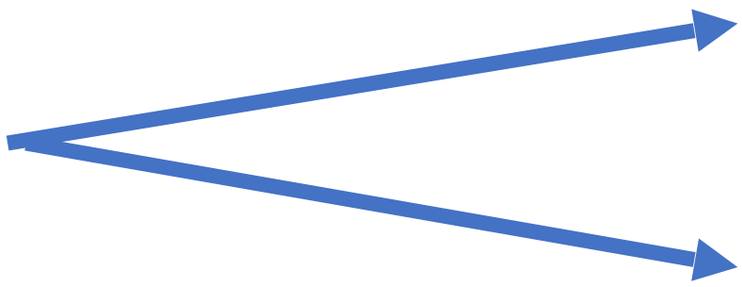
- Mrs. Patel is a 65 year old who was recently diagnosed with kidney cancer. She presents to your office. You discuss the diagnosis and treatment options. She has some questions.
  - After treatment, **what is the risk** of my cancer coming back before the Ultimate World Cruise (December 2023)?
  - **Will the risk** of my cancer coming back **change** if I get a partial nephrectomy instead of a radical nephrectomy?

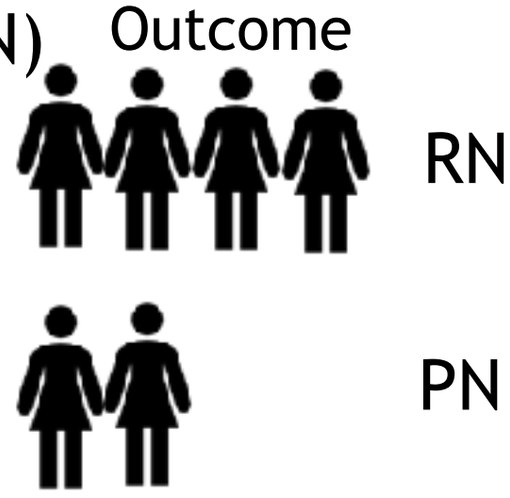
Change the risk of my cancer coming back?

- You hypothesize that type of surgery (partial vs. radical) will change her risk of cancer recurrence. How do you evaluate this hypothesis?
  - Ideally

General population of patients undergoing RN or PN

Twins of Mrs. Patel who have undergone PN or RN

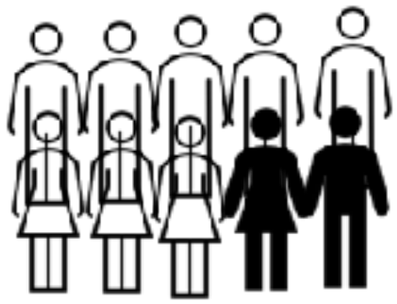




Change the risk of my cancer coming back?

- You hypothesize that type of surgery (partial vs. radical) will change her risk of cancer recurrence. How do you evaluate this hypothesis?
  - Reality

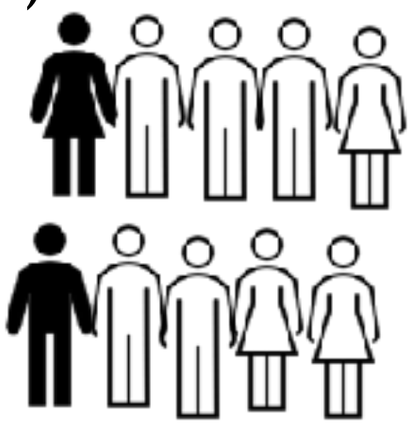
Selected population of patients undergoing RN or PN



Patients **“similar”** to Mrs. Patel who have undergone PN or RN

Surgery type (RN vs. PN)

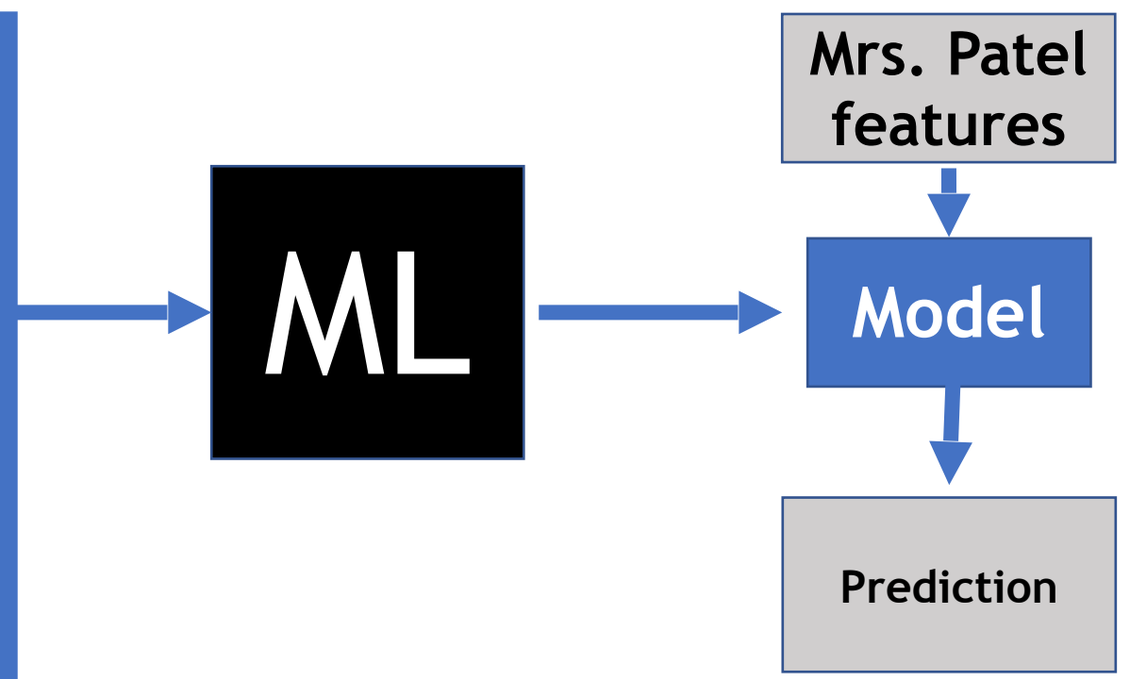
Outcome



Will my cancer come back?

- How would you the estimate of Mrs. Patel's risk of cancer recurrence?

Features X	Label Y
(age, sex, ethnicity, smoking status, body mass index, blood pressure, heart rate, …, *n*th feature)	(cancer recurrence)

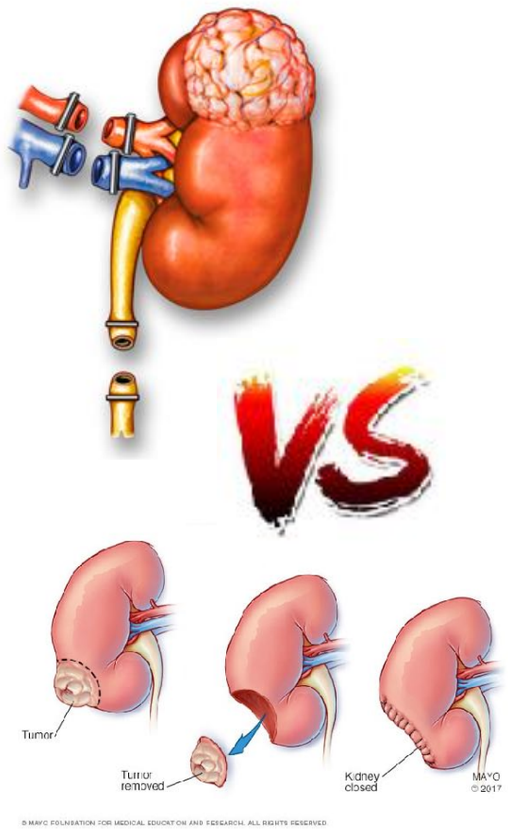


Aim of prediction is to estimate Y, given X

RCT: Radical vs. Partial Nephrectomy

- EORTC 30904

Population: 541 patients with tumors <5cm suspicious for kidney cancer



Randomized to RN vs. PN

Results

Local recurrence
RN 1/273 = 0.37%
PN 6/278 = 2.16%

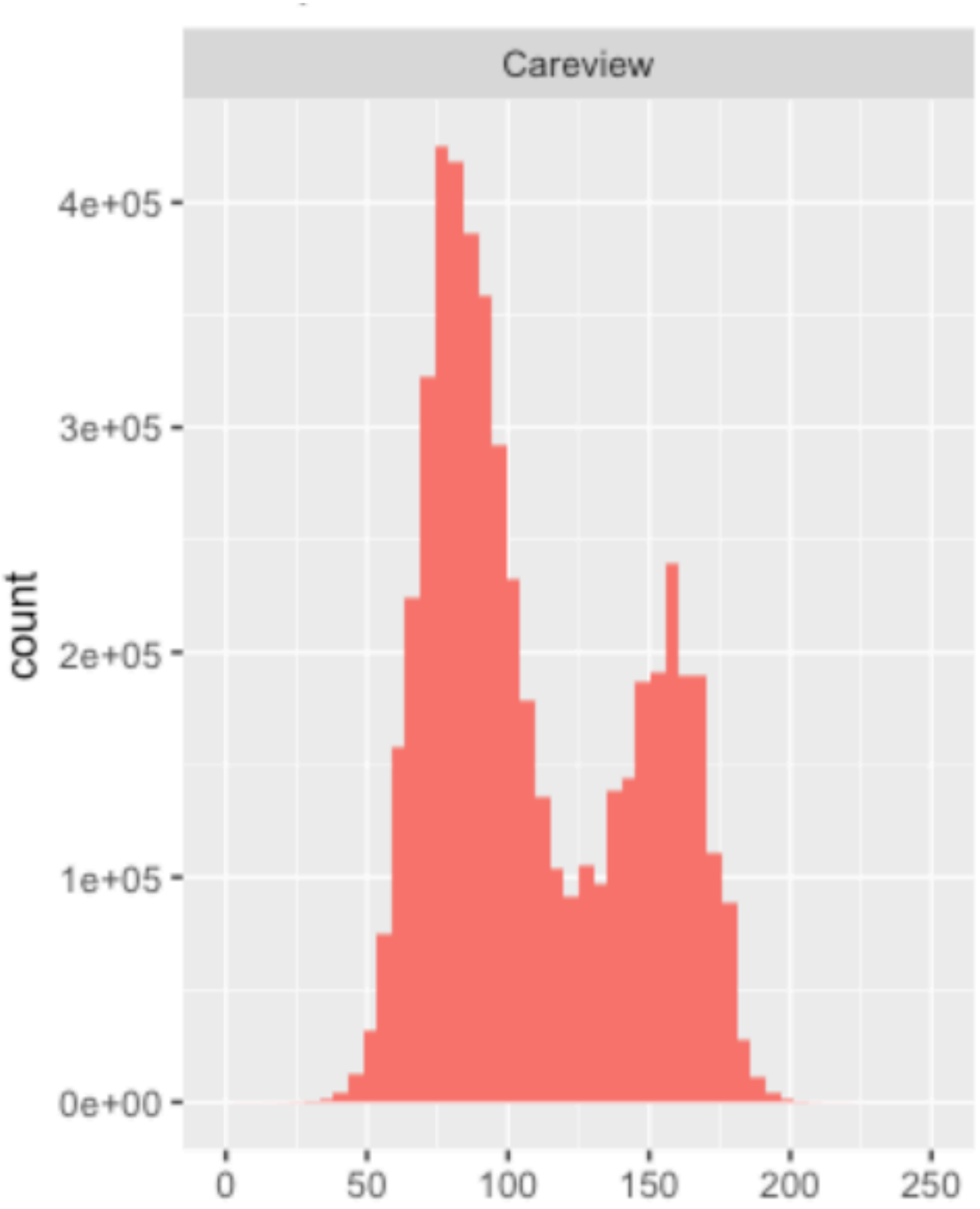
Van Poppel, Hendrik, et al. "A prospective, randomised EORTC intergroup phase 3 study comparing the oncologic outcome of elective nephron-sparing surgery and radical nephrectomy for low-stage renal cell carcinoma." *European urology* 59.4 (2011): 543-552.
https://www.fairbanksurology.com/robotic-radical-nephrectomy
https://www.mayoclinic.org/tests-procedures/nephrectomy/multimedia/img-20332175

MIMIC-IV
https://physionet.org/content/mimiciv/

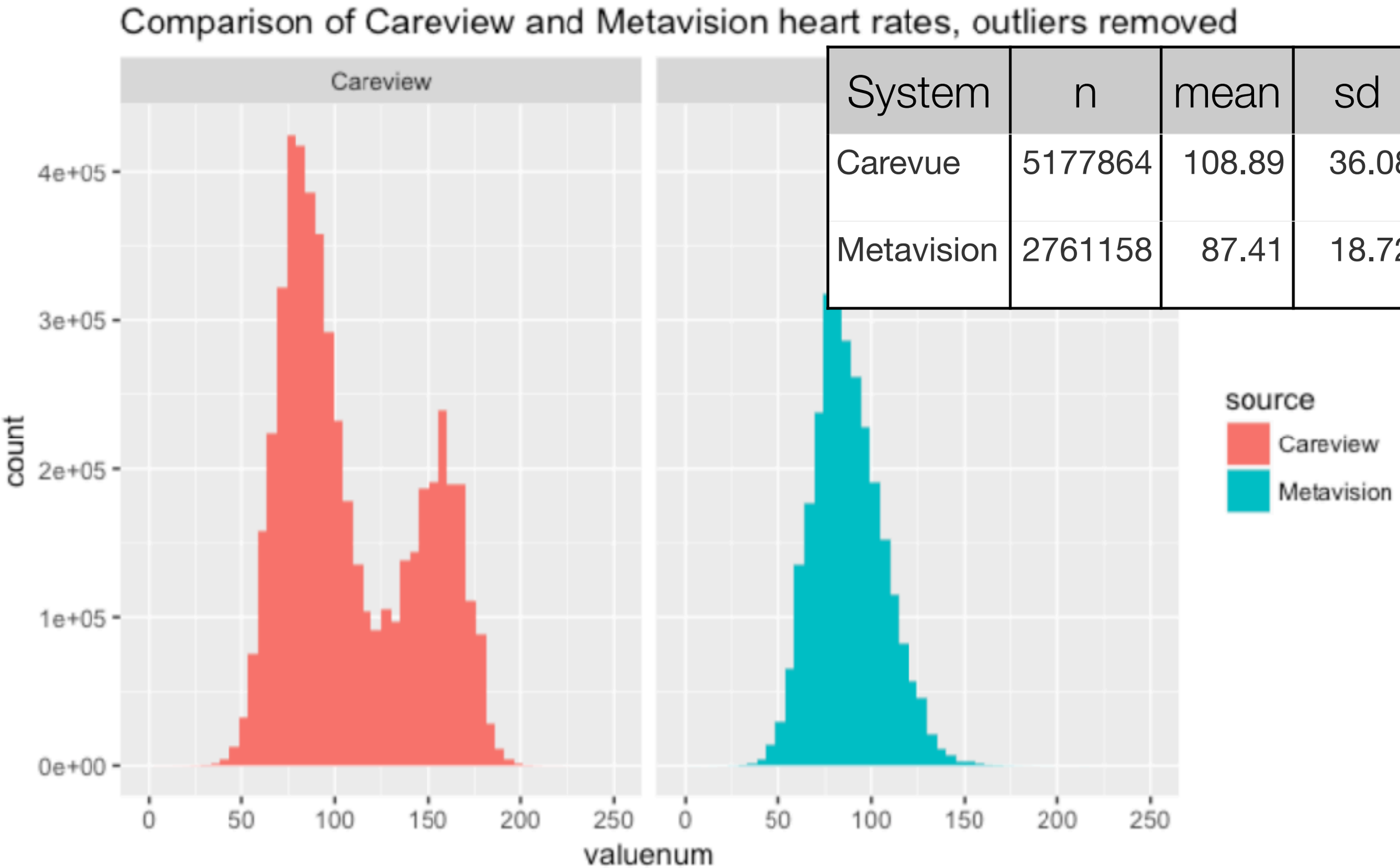
- **Data Dictionaries**
  - d_icd_diagnoses
  - d_icd_procedures
  - d_hcpcs [CPT+]
  - d_items
  - d_labitems
- **Hospitalization-related**
  - patients
  - admissions
  - hcpcsevents
  - diagnoses_icd
  - procedures_icd
  - drgcodes
  - omr
  - emar [medication administration]
    - emar_detail
  - labevents
  - microbiologyevents
  - poe
    - poe_detail
  - prescriptions
    - pharmacy
  - services
  - transfers
- **ICU-related**
  - icustays
  - chartevents
  - inputevents
    - ingredientevents
  - outputevents
  - procedureevents
  - datetimeevents

Understanding clinical data

- Consider the distribution of heart rates in the MIMIC-III chart (as recorded in Carevue)

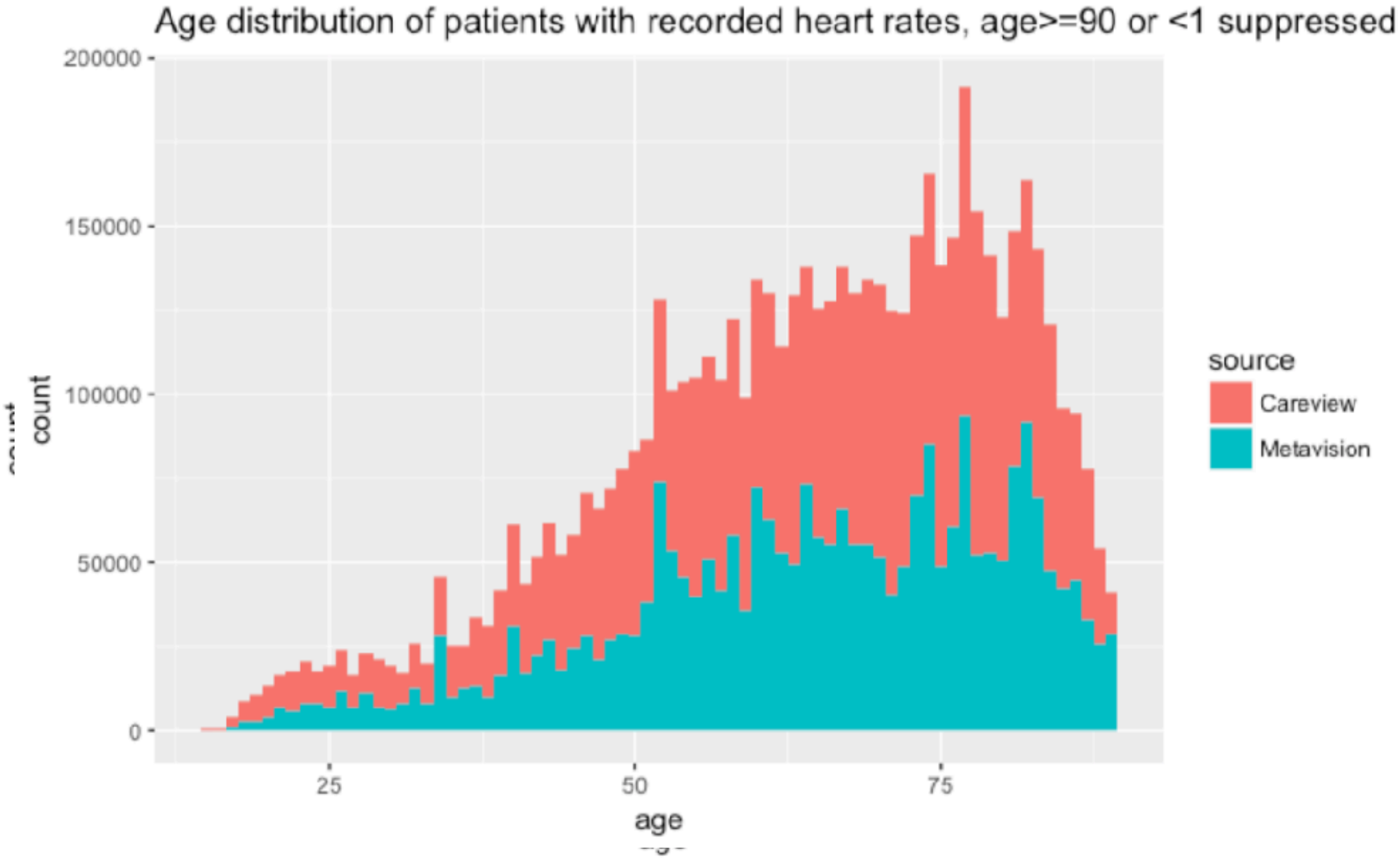


Data from https://mimic.physionet.org

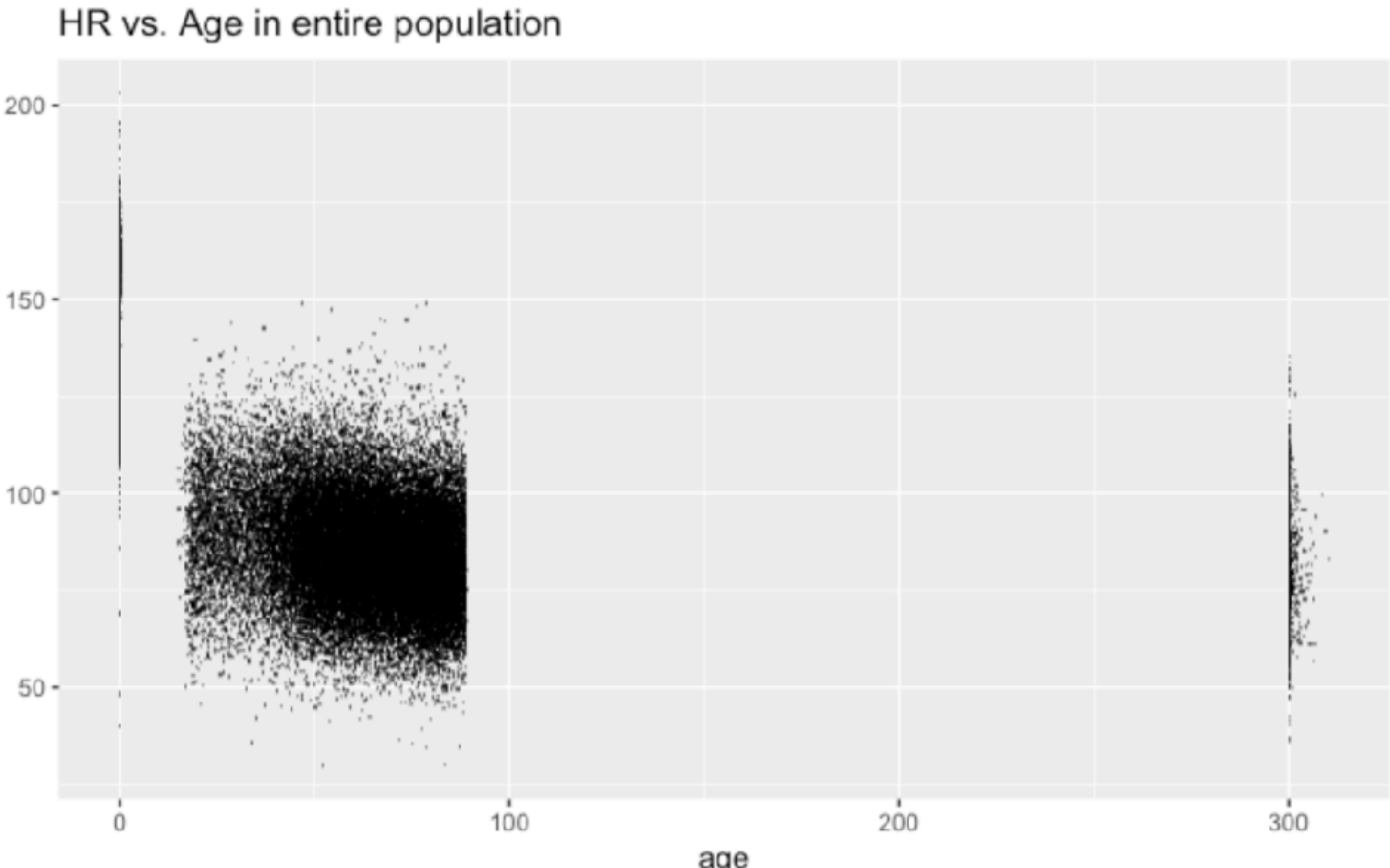


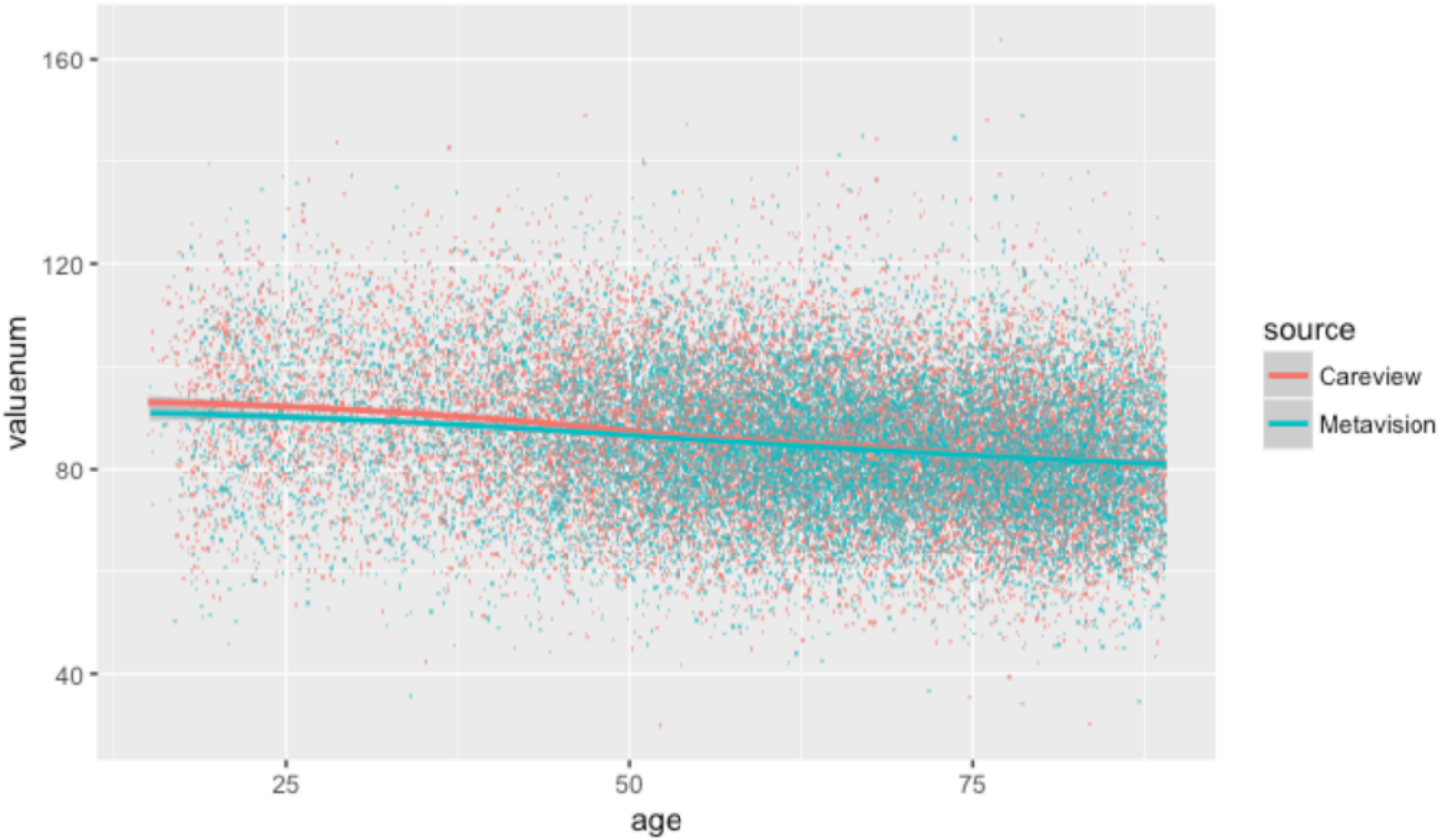
System	n	mean	sd
Carevue	5177864	108.89	36.08
Metavision	2761158	87.41	18.72

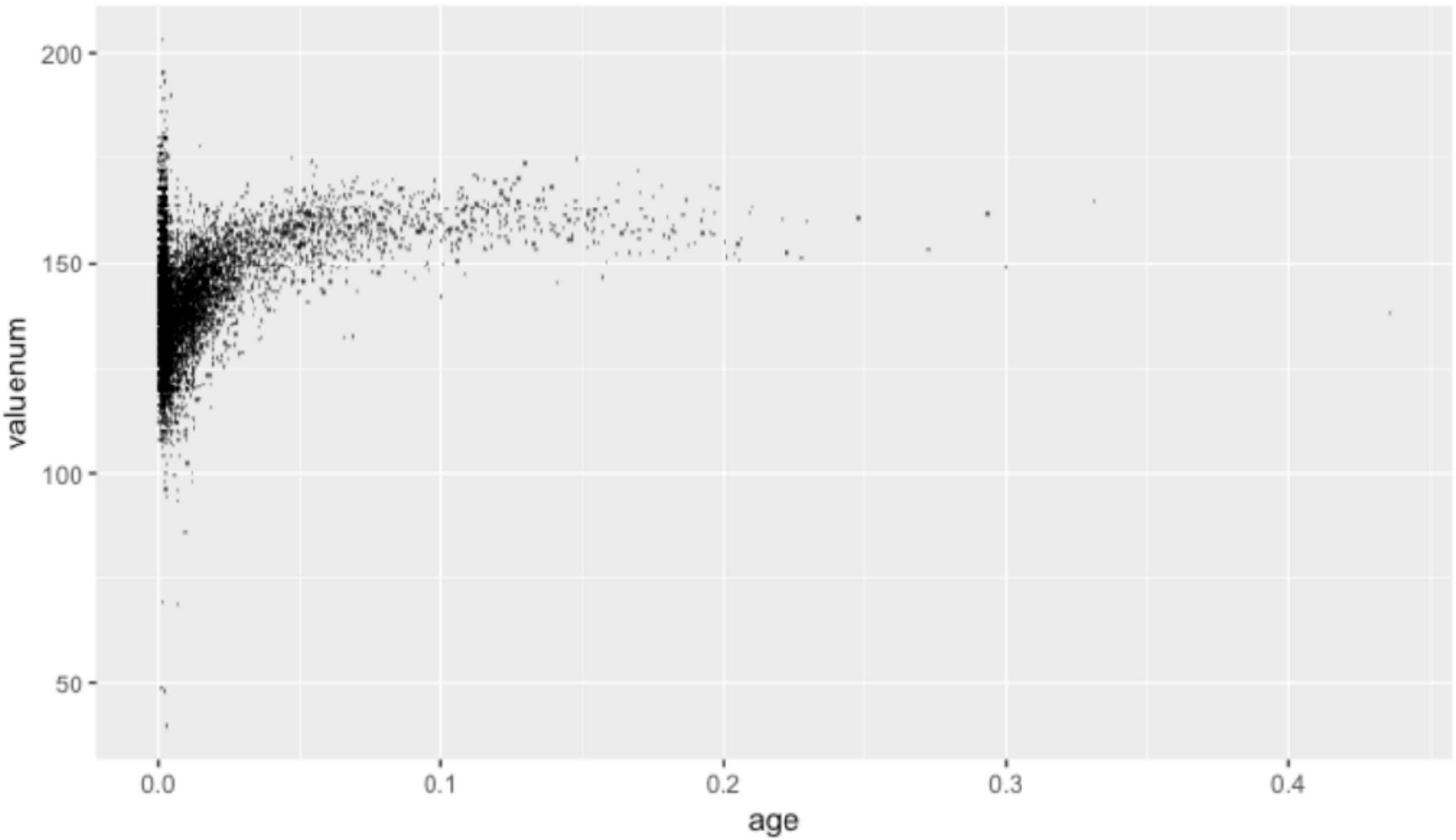

Age distribution of patients with recorded heart rates, age>=90 suppressed
Careview
Metavision
1500000
1000000
500000
0
count
0
25
50
75
age
source
Careview
Metavision

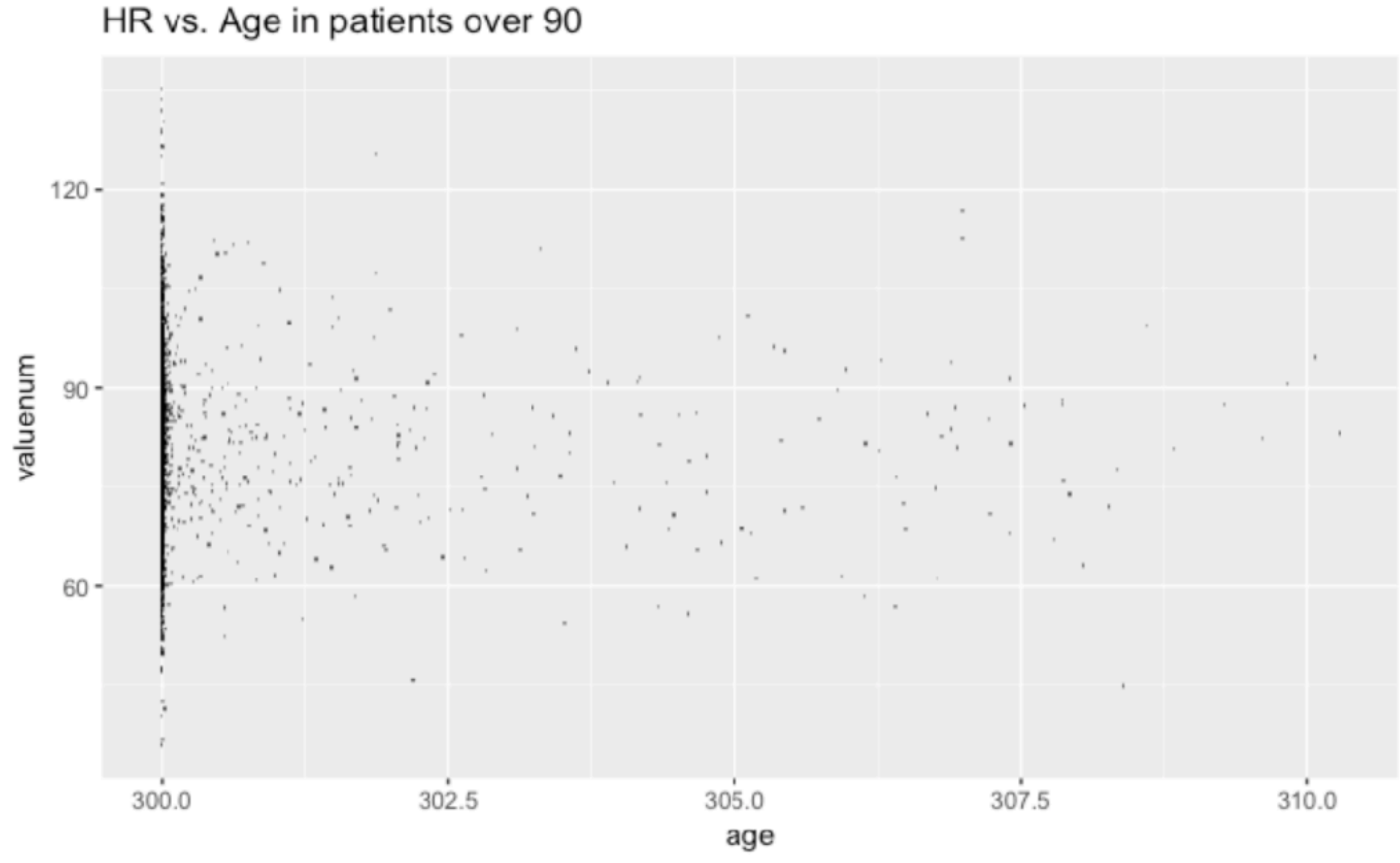

Age distribution of patients with recorded heart rates, age>=90 or <1 suppressed
200000
150000
100000
50000
0
count
25
50
75
age
source
Careview
Metavision

Is Age a confounder for Heart Rate?

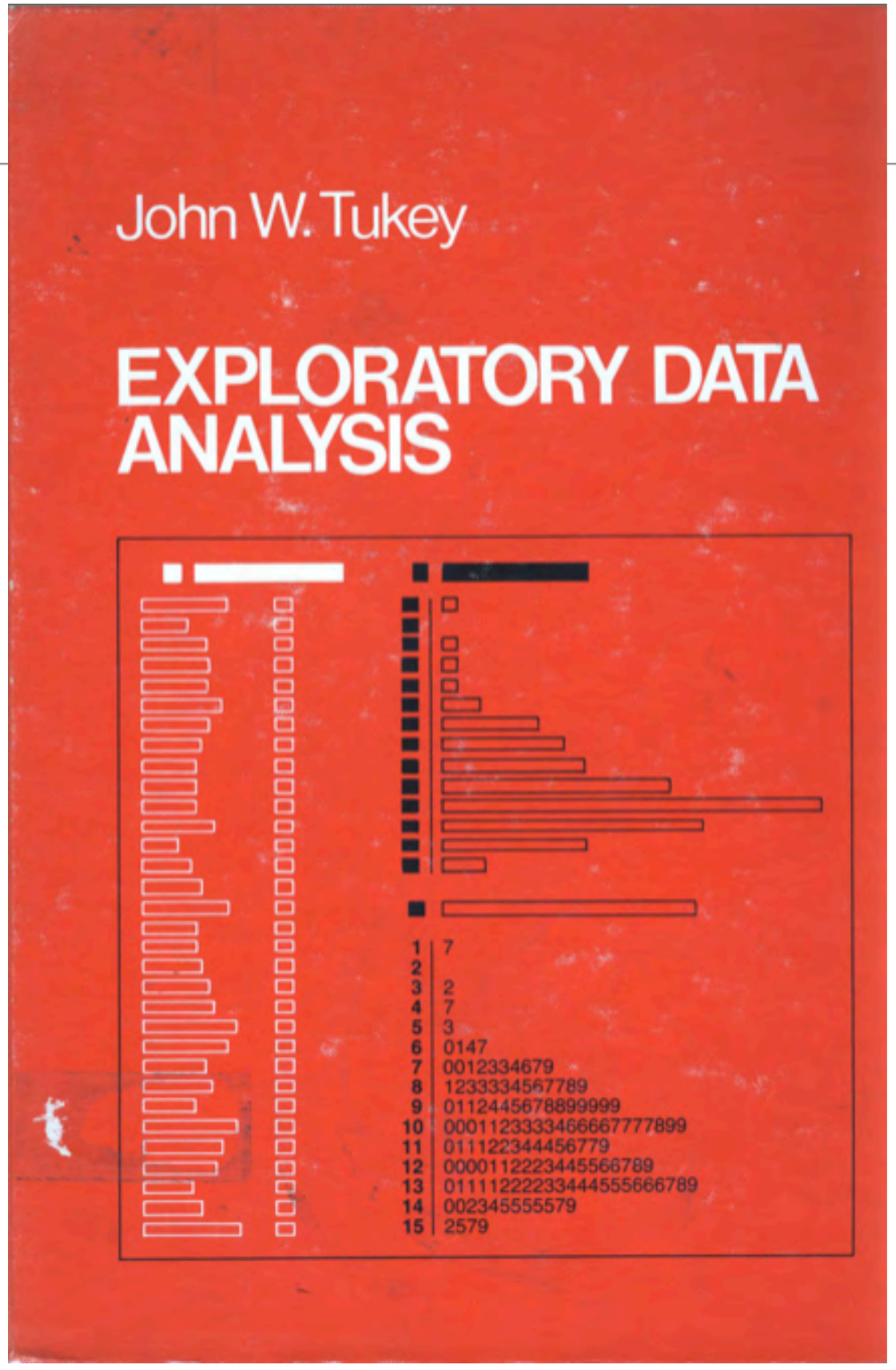


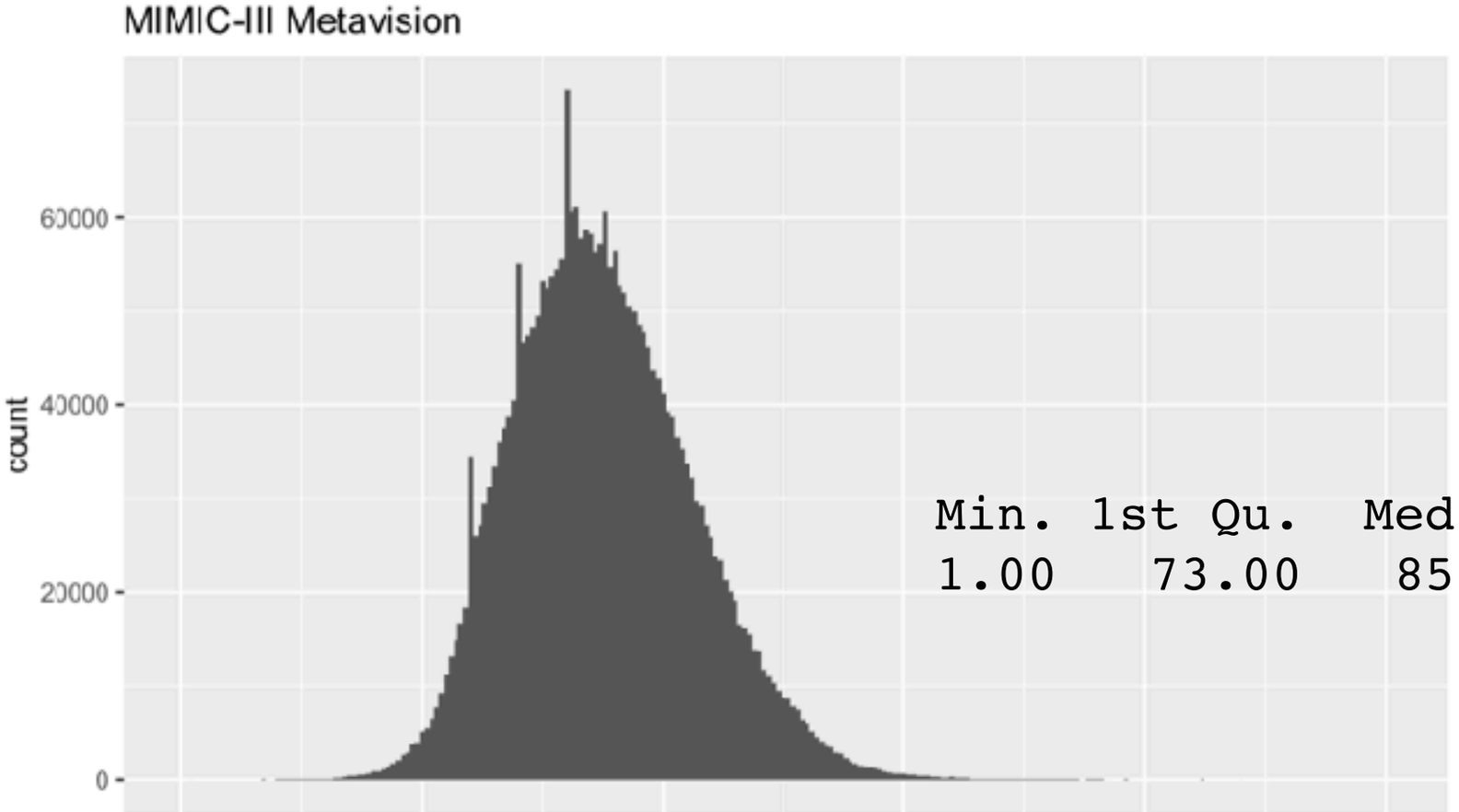

HR vs. Age in adults, smoothed
160
120
80
40
value_numeric
25
50
75
age
source
Careview
Metavision


HR vs. Age in neonates
200
150
100
50
valuenum
0.0
0.1
0.2
0.3
0.4
age


HR vs. Age in patients over 90
120
90
60
valuenum
300.0
302.5
305.0
307.5
310.0
age

EDA

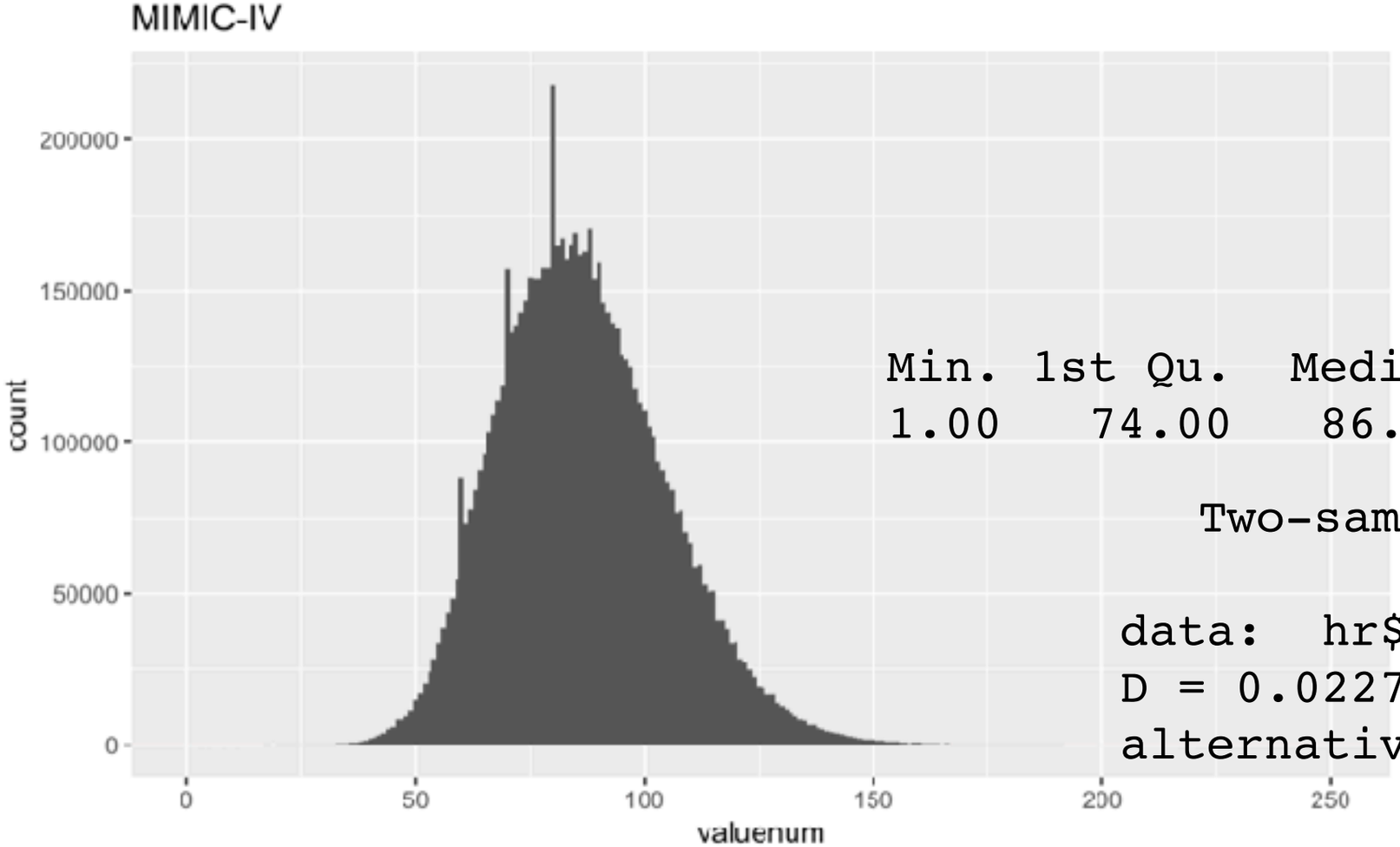

John W. Tukey
EXPLORATORY DATA ANALYSIS
1 7
2
3 2
4 7
5 3
6 0147
7 0012334679
8 1233334567789
9 0112445678899999
10 000112333346666777899
11 011122344456779
12 0000112223445566789
13 011112222334445556667789
14 002345555579
15 2579



MIMIC-IV and MIMIC-III Metavision Heart Rates are very similar

N=2,761,158

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1.00   73.00   85.00   86.51   98.00  250.00
```



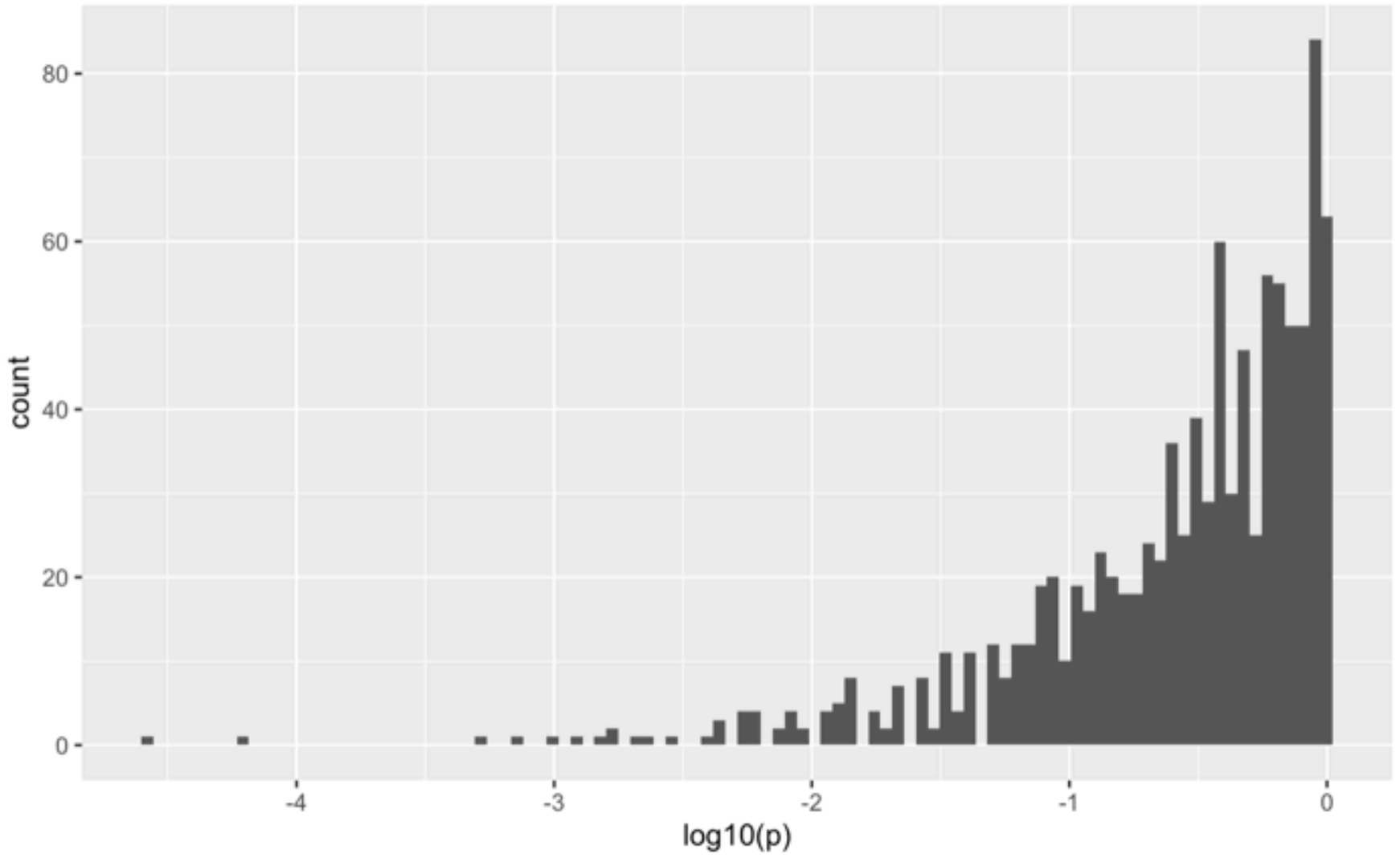
N=7,625,316

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1.00   74.00   86.00   87.41   99.00  250.00
```

```
        Two-sample Kolmogorov-Smirnov test

data:  hr$valuenum and hr3$VALUENUM
D = 0.02271, p-value < 2.2e-16
alternative hypothesis: two-sided
```

Be careful with statistics!

- 1000 times:
- Sampled 1000 data points from the MIMIC-III and -IV HR distributions
- Checked K-S test to see if we could reject the null (that they came from the same distribution)
- Look at the distribution of p-values



Types of Data

- Demographics
  - Age, sex, socio-economic status, insurance type, language, religion, living situation, family structure, location, work, …
- Vital signs
  - Weight, height, pulse, respiration rate, body temperature, …
- Medications
  - Prescriptions, over-the-counter drugs, illegal drugs, alcohol, …
  - Medication reconciliation
- Laboratory
  - Components of blood, urine, stool, saliva, spinal fluid (CSF), ascitic fluid, joint fluid, bone marrow, lung, …
- Images
  - MRI, CT, PET, X-ray, ultrasound, retinoscopy, endoscopy, photographs, …
- Signals
  - ECG, EMG, EEG, EMG, continuous blood pressure, …
- Genetics
  - Sequences, transcriptome, single-cell, proteome, …

Types of Data (continued)

- Pathology
  - Qualitative and quantitative examination of any body tissue, e.g., biopsy samples, surgical “scraps”
  - Cell-level measurements, e.g., cell-surface antigens Microbiology — organisms grown, typically from cultures
  - Testing sensitivity to various antibiotics, at various dilutionsInput/Output (fluids)
- Notes
  - Discharge summary
  - Attending and/or Resident
  - Nurse
  - Specialist
    - Radiology, Pathology, ECG, Nutrition, Respiratory, Social work, …
  - Consultant
  - Referring physician
  - Emergency Department

Types of Data (continued)

- Billing
  - Diagnoses (ICD-{9, 10})
  - Procedures (CPT and ICD)
  - Diagnosis Related Groups (DRG) [~ abstraction of ICD]
- Administrative
  - Service
  - Transfers

Types of Data (continued)

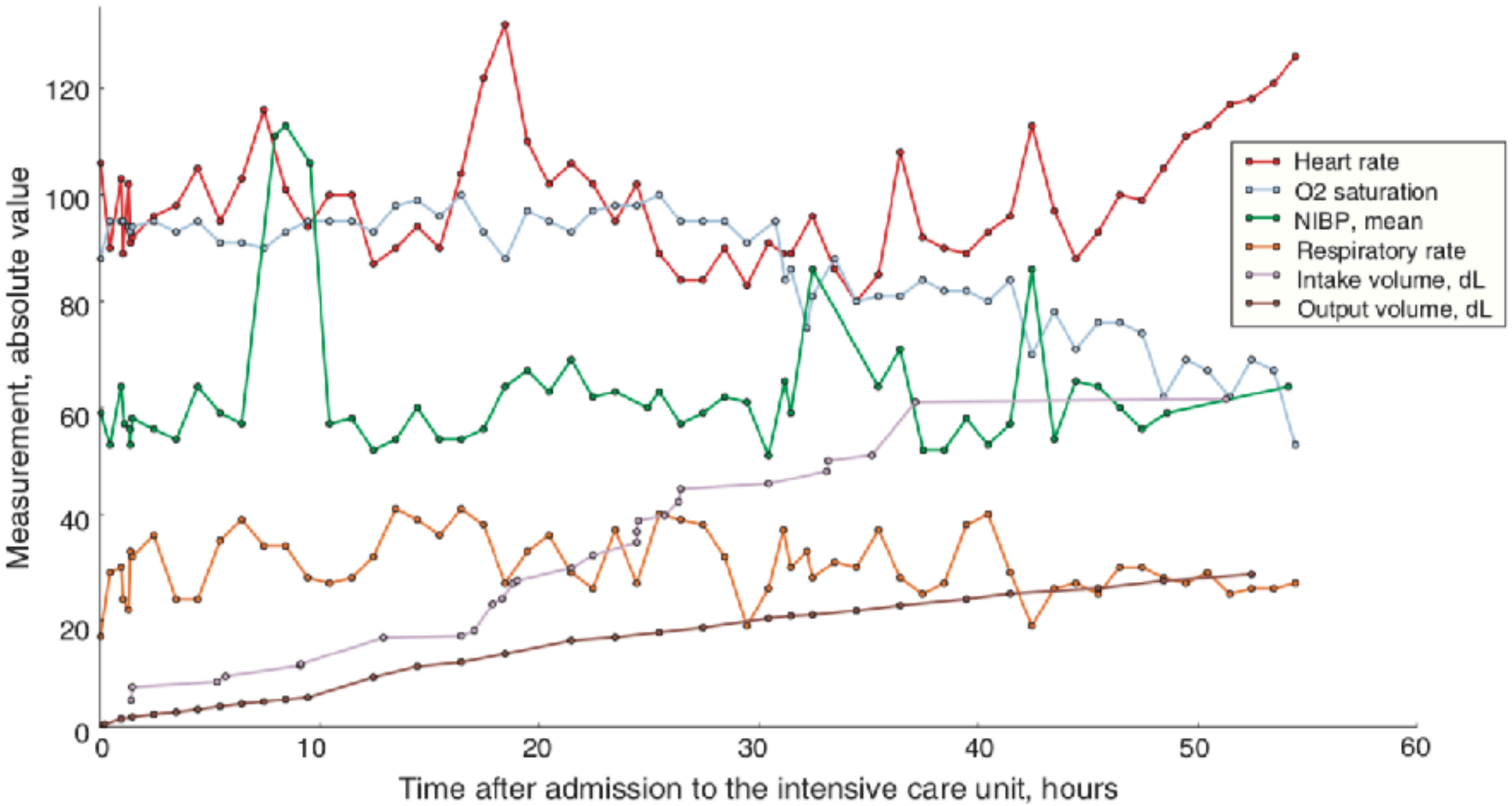
- Quantified Self
  - Activity
    - Steps
    - Elevation change
    - Workouts
  - Vitals
    - Heart rate
    - Respiration rate
    - Temperature
    - Blood pressure
    - Weight
- Diet
- Blood sugar
- Allergies
- Mindfulness
- Mood
- Sleep
- Pain
- Sex

- “N-of-1 experiments”
- Growing availability of home health measurements

Code status	Full code				Comfort measures	
GCS: Verbal	Oriented	Oriented	Oriented	Confused	Confused	Incomprehensible sounds
GCS: Motor	Obeys commands	Obeys commands	Obeys commands	Obeys commands	Obeys commands	Flex-withdraws
GCS: Eye	Spontaneously	Spontaneously	Spontaneously	To speech	To speech	None
Platelet, K/uL	48 53	46			45	
Creatinine, mg/dL	0.7	0.7			0.8	
White blood cell, K/uL	9.1 12.4	16.8			23.2	
Neutrophil, %	37					

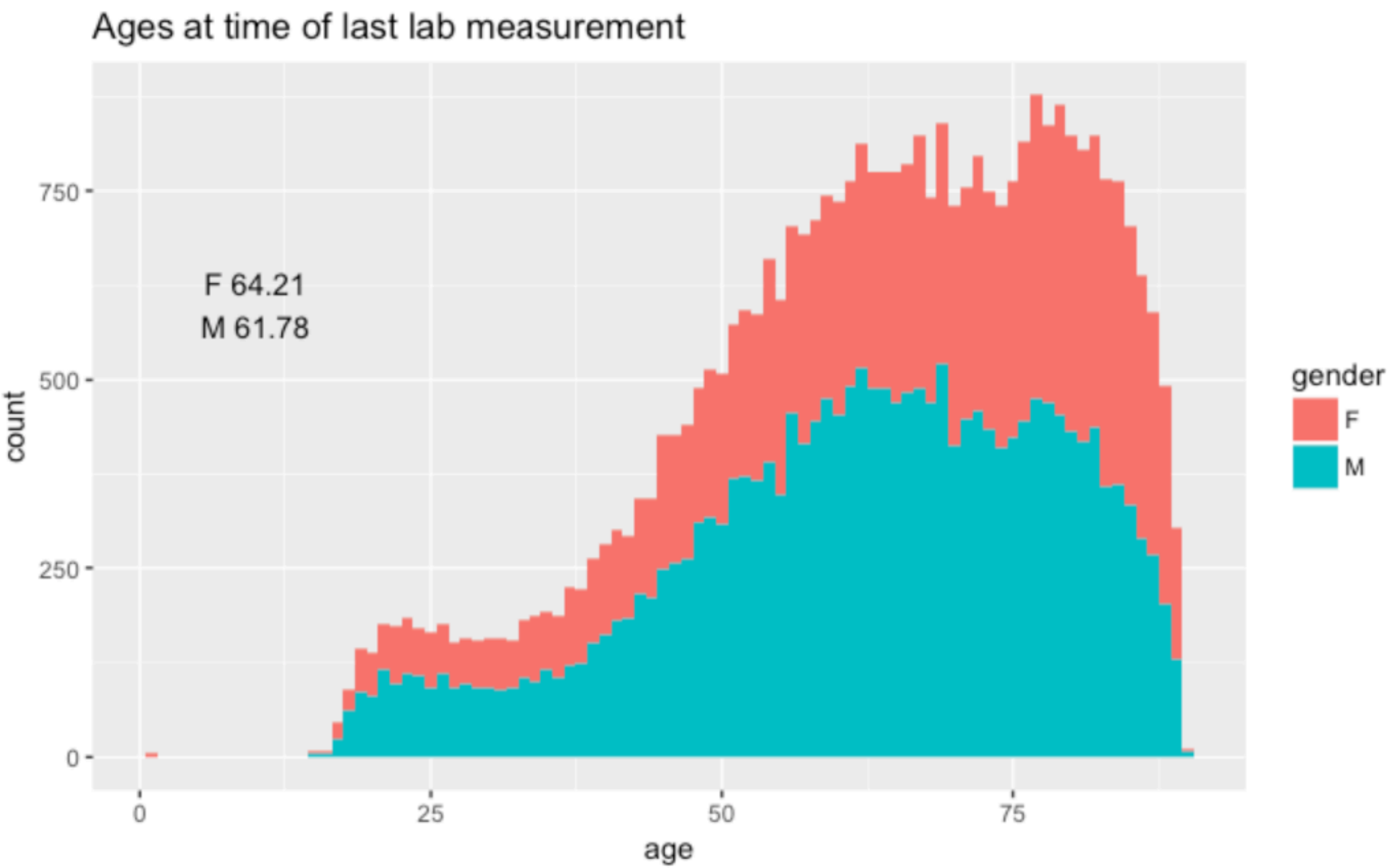
Morphine Sulfate
Vancomycin (1 dose)
Piperacillin (1 dose)
NaCl 0.9%
Amiodarone
Dextrose 5%

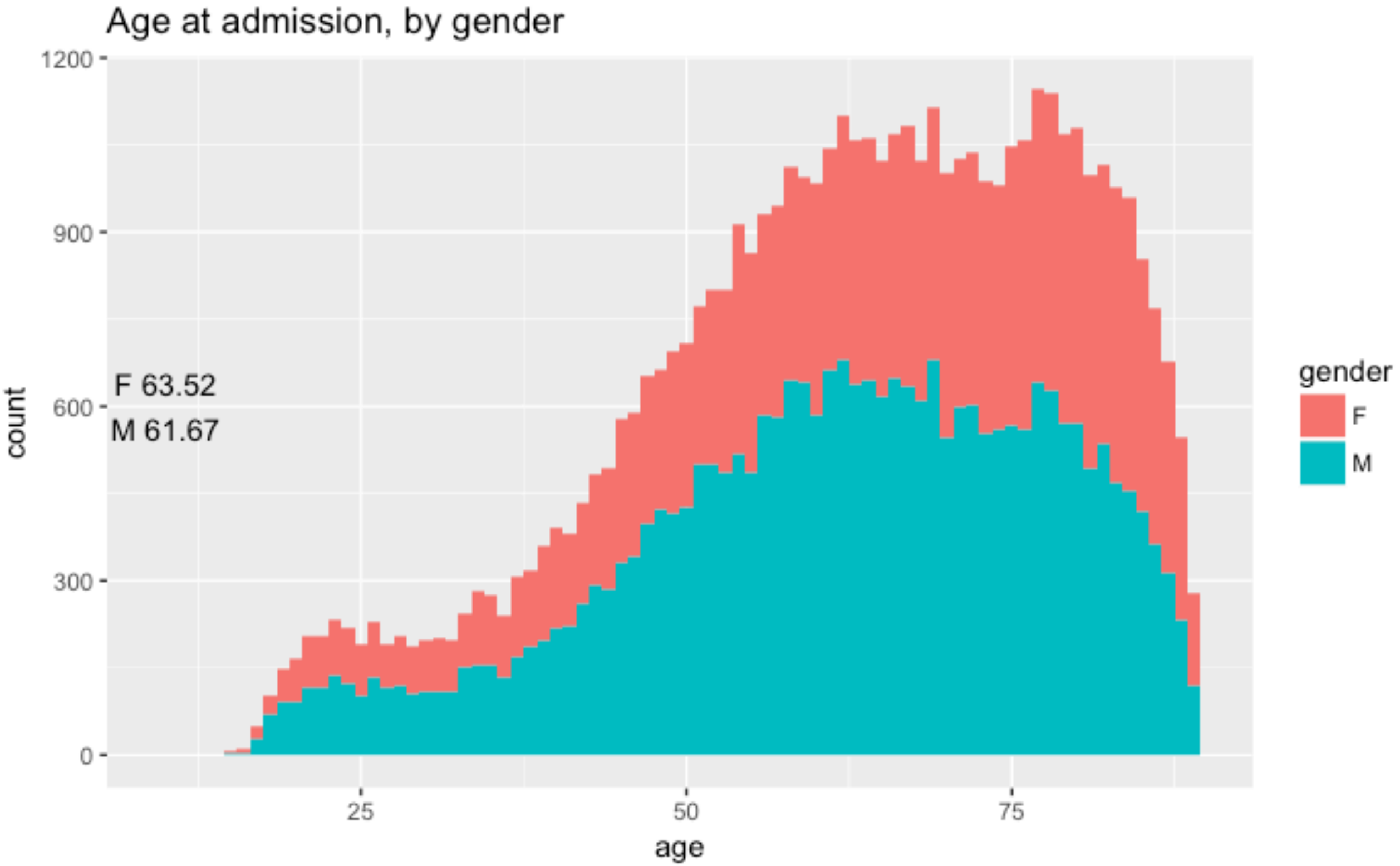
10.0 mL/hour — 10.0mL/hour — 10.0mL/hour
1 mg/min — 0.5mg/min — 0.5mg/min
50mL/hour — 25mL/hour — 25mL/hour

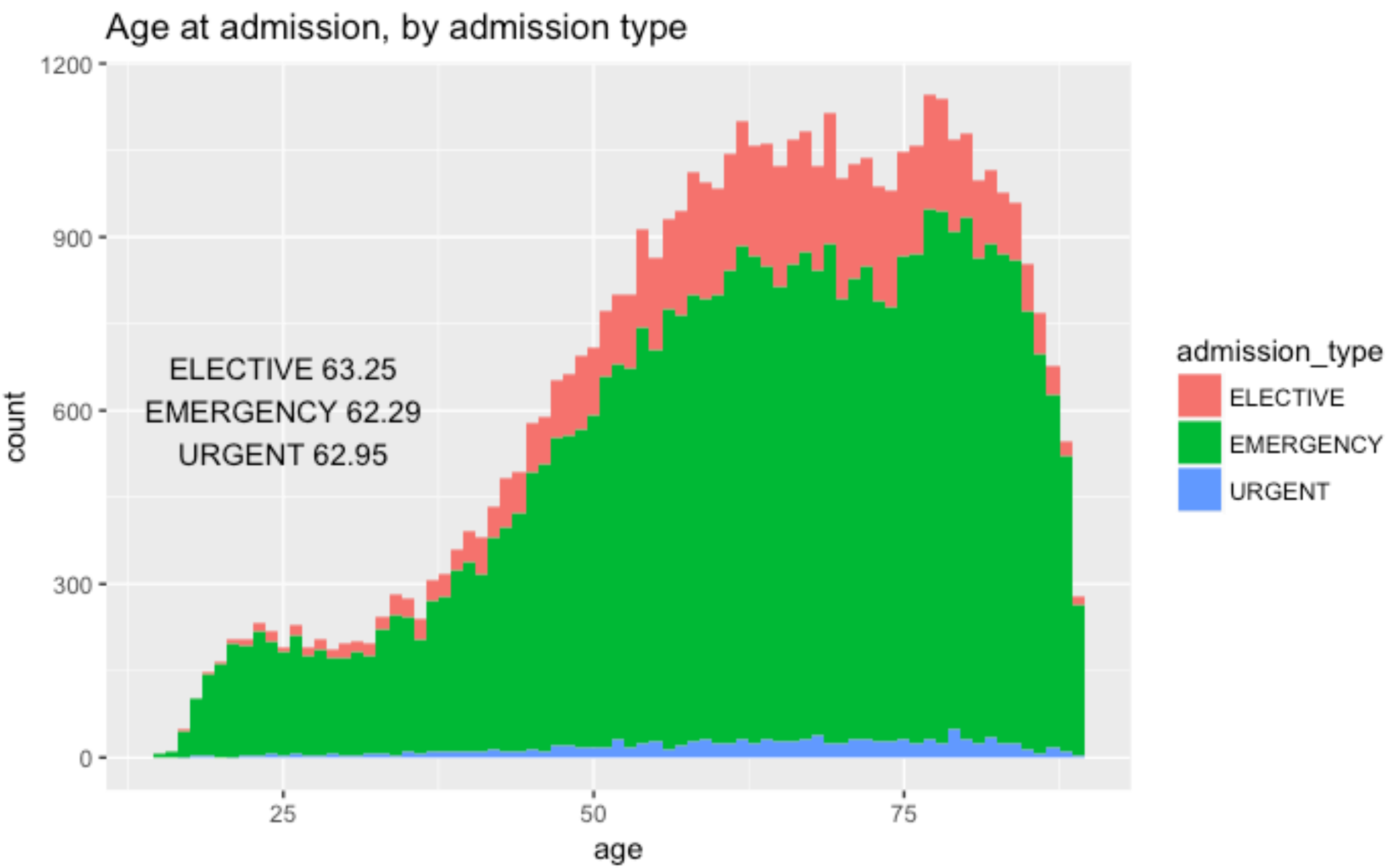


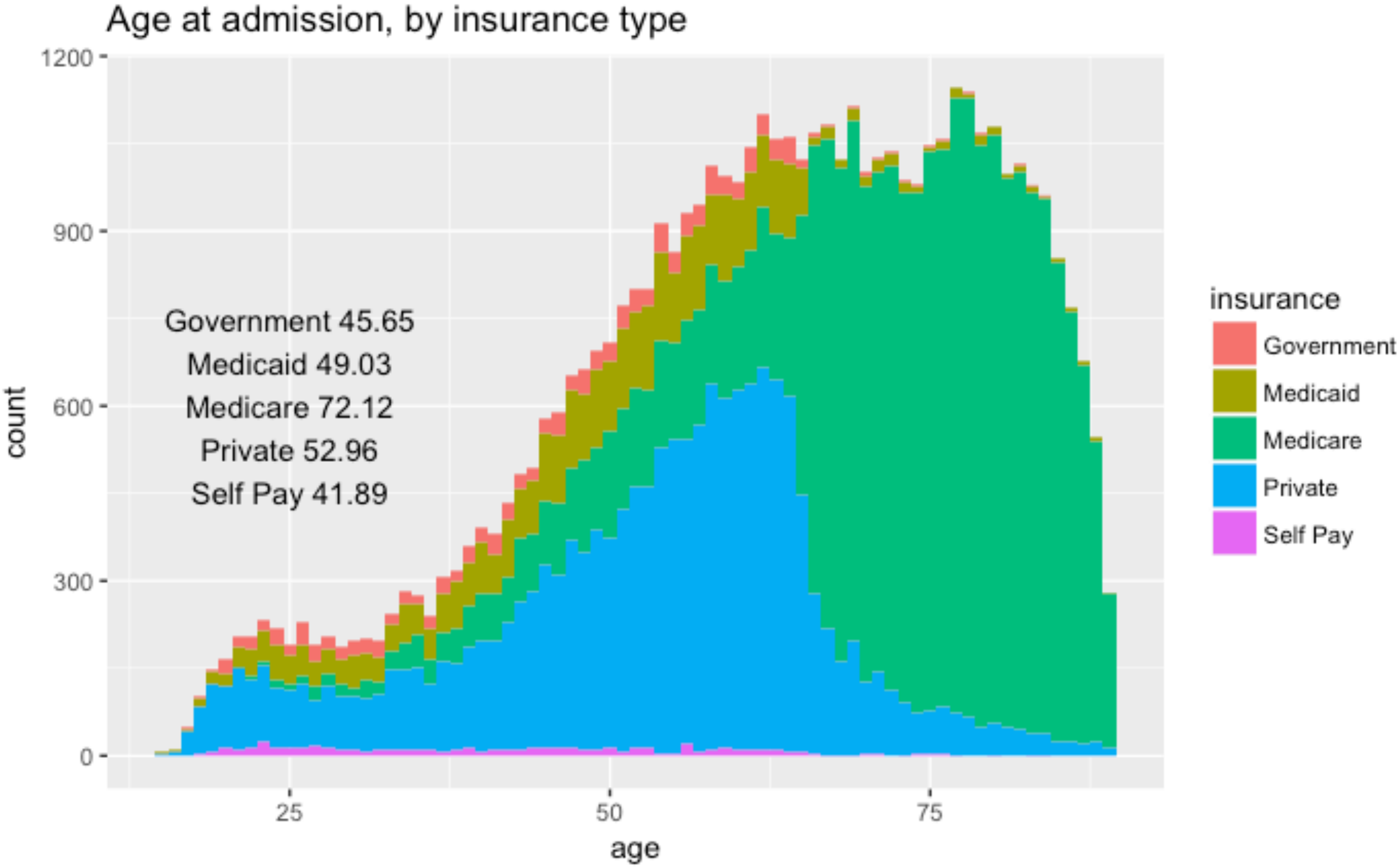
Demographics

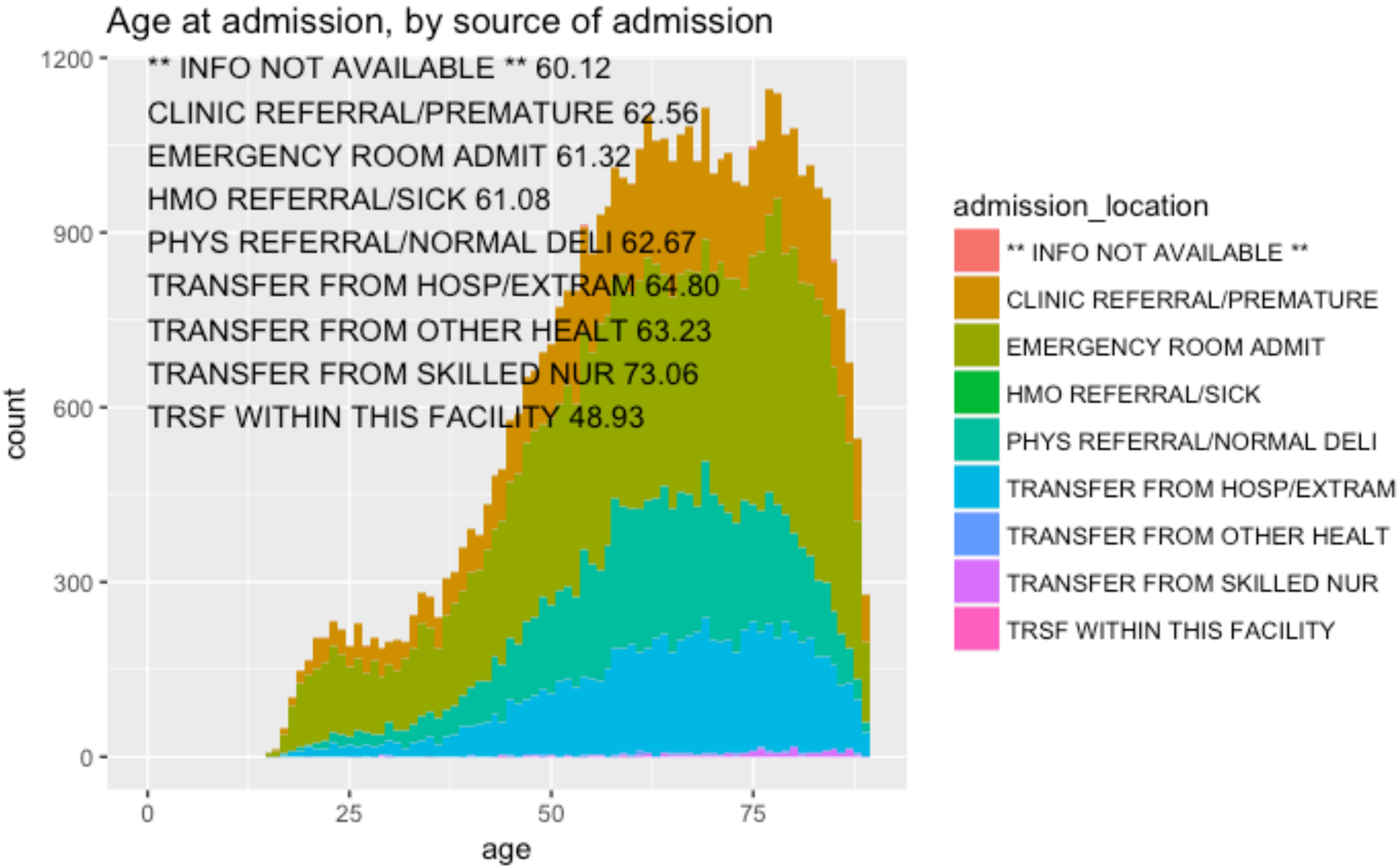
- Consider how the age distribution changes by
  - gender
  - type of admission
  - type of insurance
  - source of admission
  - whether they die during the admission
  - native language
  - ethnicity
  - marital status

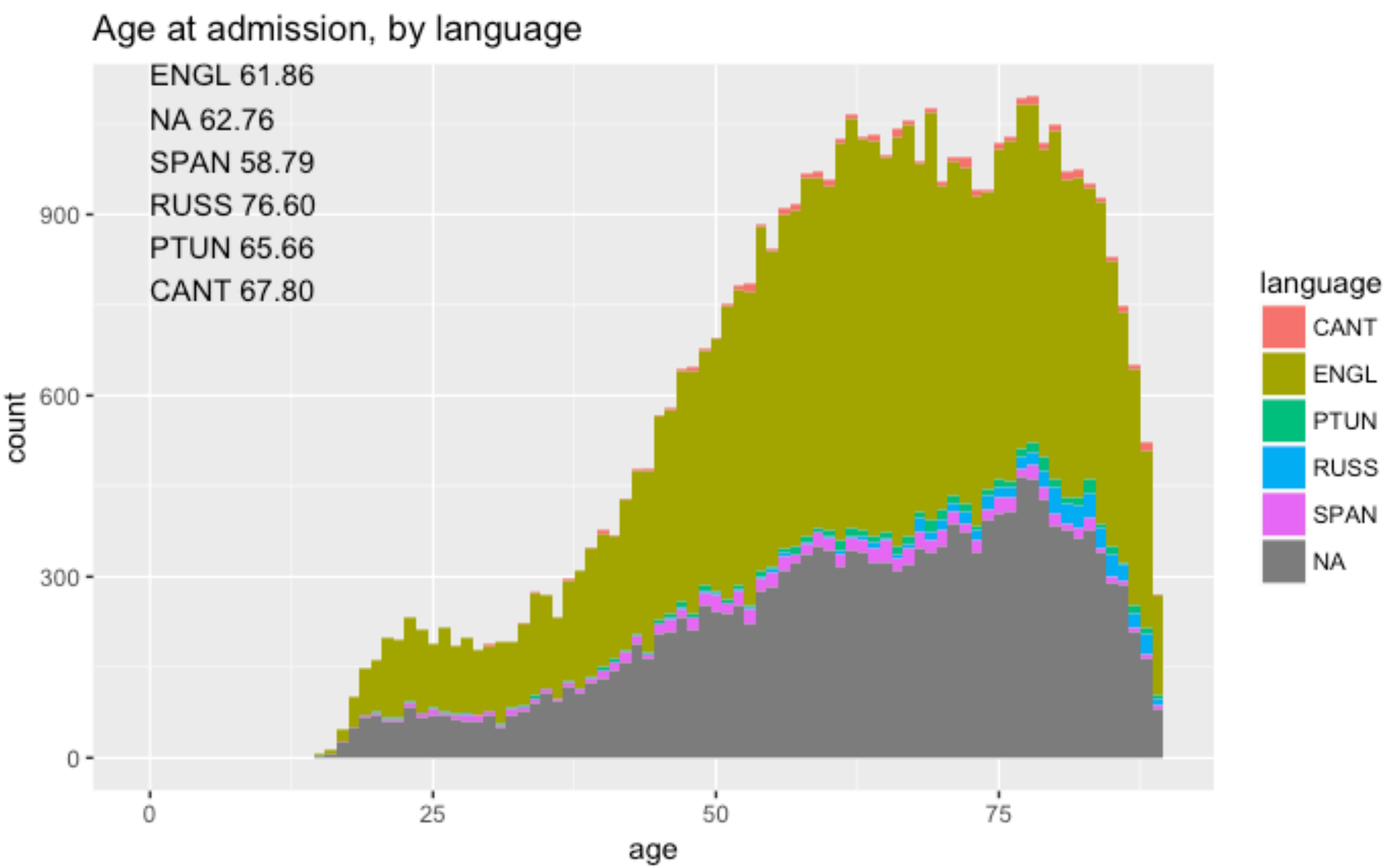

Ages at time of last lab measurement
F 64.21
M 61.78
750
500
250
0
count
0
25
50
75
age
gender
F
M


Age at admission, by gender
1200
900
600
300
0
count
F 63.52
M 61.67
25
50
75
age
gender
F
M

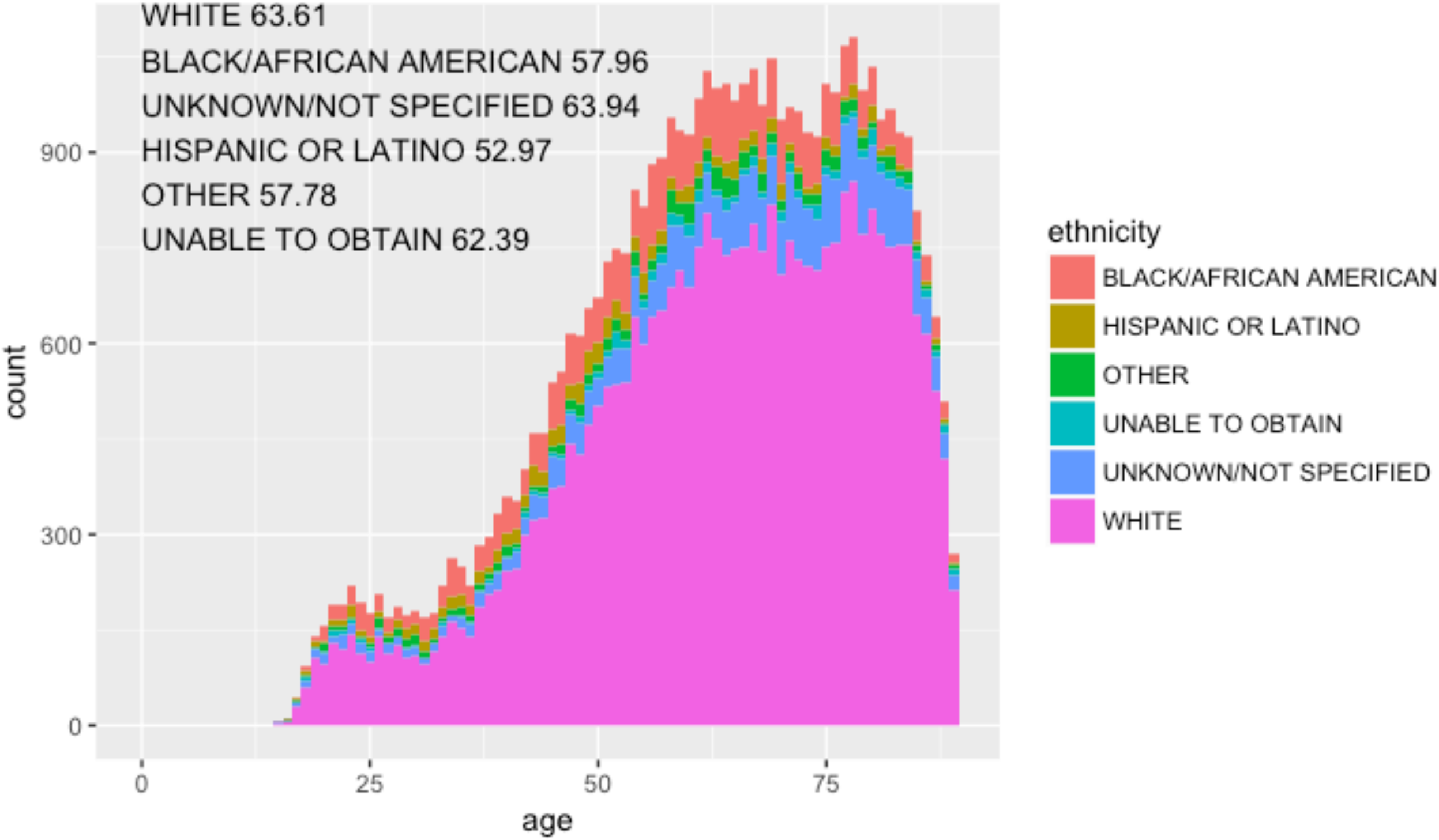

Age at admission, by admission type
1200
900
600
300
0
count
ELECTIVE 63.25
EMERGENCY 62.29
URGENT 62.95
25
50
75
age
admission_type
ELECTIVE
EMERGENCY
URGENT

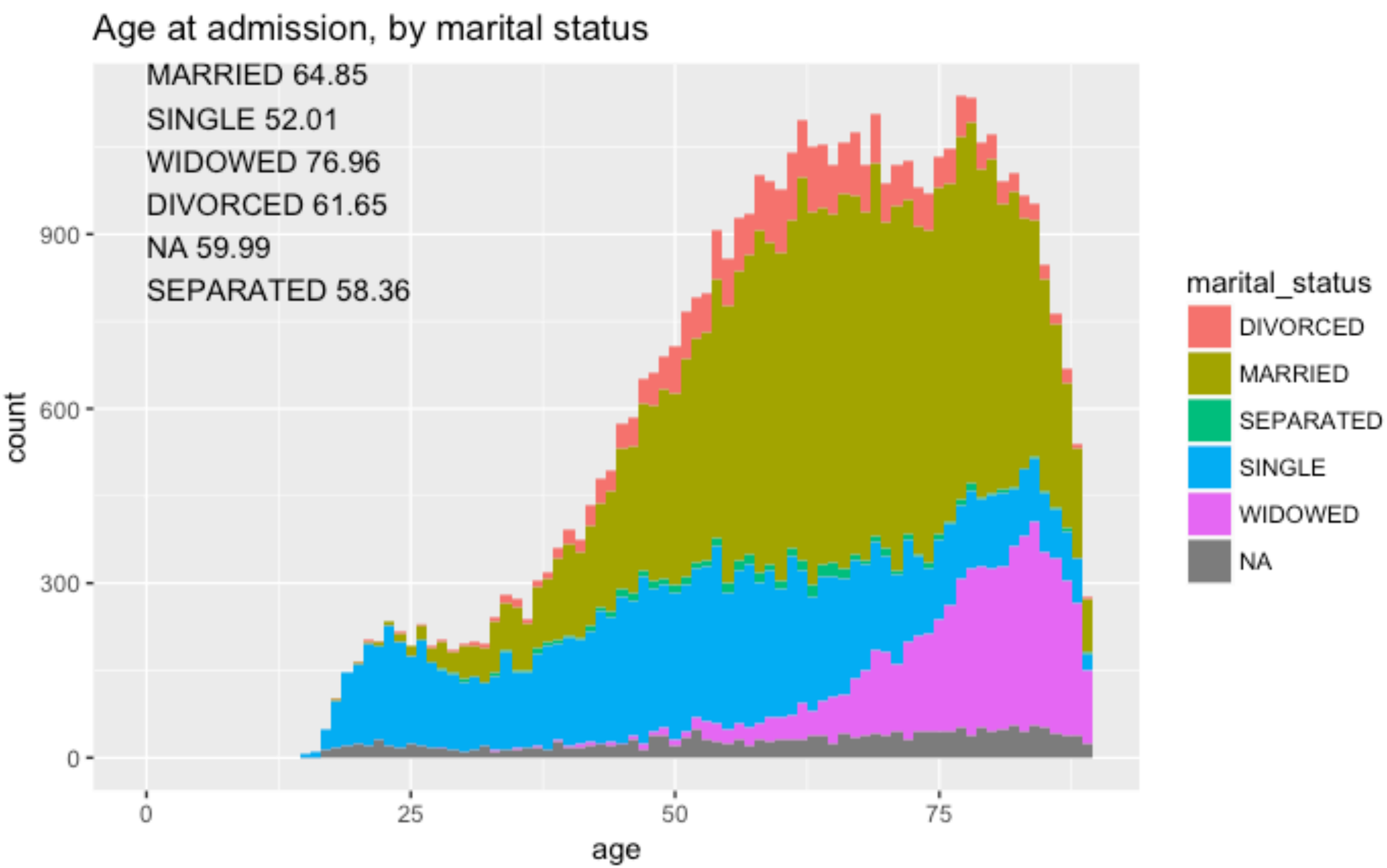

Age at admission, by insurance type
1200
900
600
300
0
count
Government 45.65
Medicaid 49.03
Medicare 72.12
Private 52.96
Self Pay 41.89
25
50
75
age
insurance
Government
Medicaid
Medicare
Private
Self Pay


Age at admission, by source of admission
** INFO NOT AVAILABLE ** 60.12
CLINIC REFERRAL/PREMATURE 62.56
EMERGENCY ROOM ADMIT 61.32
HMO REFERRAL/SICK 61.08
PHYS REFERRAL/NORMAL DELI 62.67
TRANSFER FROM HOSP/EXTRAM 64.80
TRANSFER FROM OTHER HEALT 63.23
TRANSFER FROM SKILLED NUR 73.06
TRSF WITHIN THIS FACILITY 48.93
count
1200
900
600
300
0
0
25
50
75
age
admission_location
** INFO NOT AVAILABLE **
CLINIC REFERRAL/PREMATURE
EMERGENCY ROOM ADMIT
HMO REFERRAL/SICK
PHYS REFERRAL/NORMAL DELI
TRANSFER FROM HOSP/EXTRAM
TRANSFER FROM OTHER HEALT
TRANSFER FROM SKILLED NUR
TRSF WITHIN THIS FACILITY


Age at admission, by language
ENGL 61.86
NA 62.76
SPAN 58.79
RUSS 76.60
PTUN 65.66
CANT 67.80
900
600
300
0
count
0
25
50
75
age
language
CANT
ENGL
PTUN
RUSS
SPAN
NA

Age at admission, by ethnicity



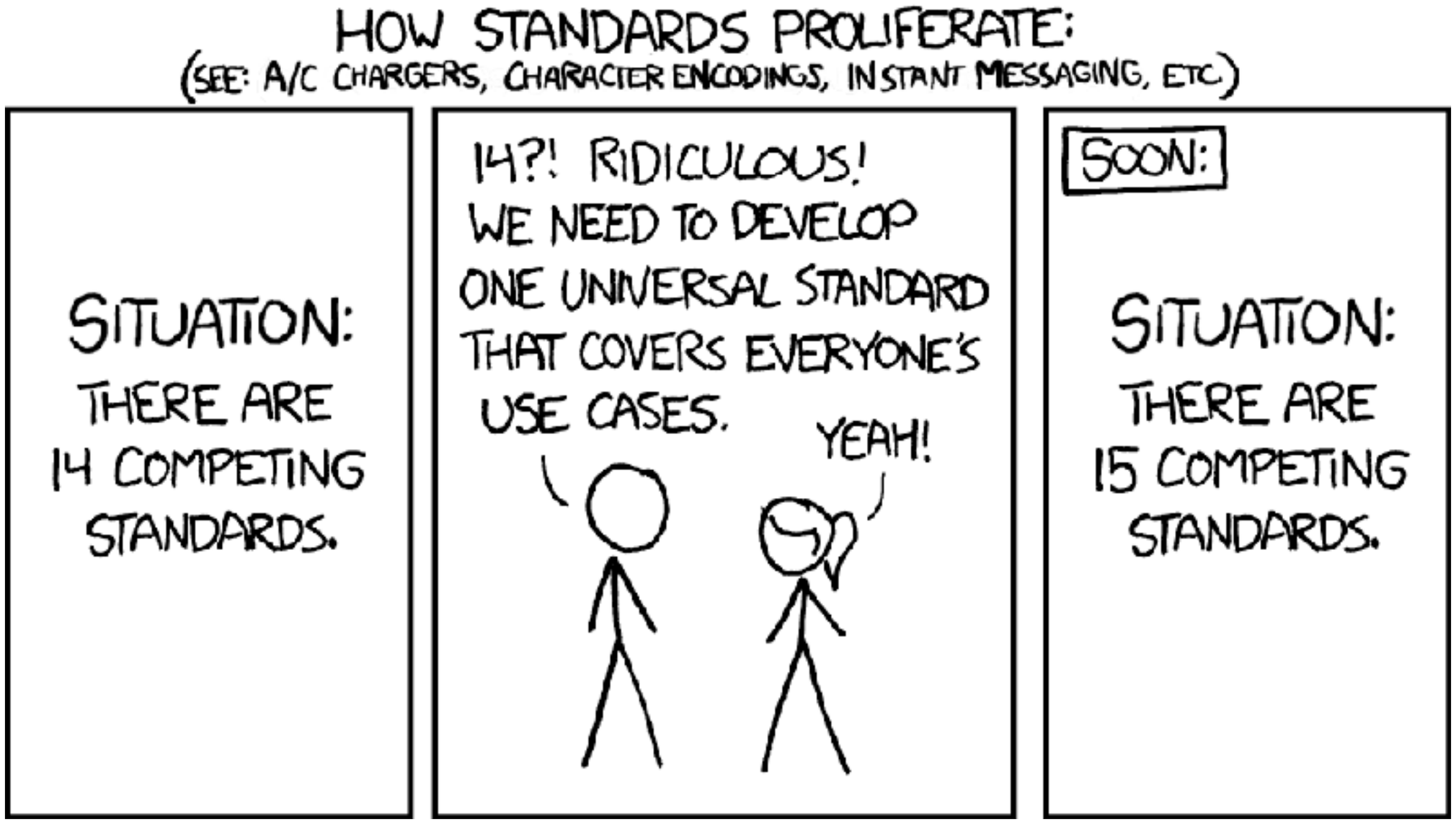

Age at admission, by marital status
MARRIED 64.85
SINGLE 52.01
WIDOWED 76.96
DIVORCED 61.65
NA 59.99
SEPARATED 58.36
count
900
600
300
0
0
25
50
75
age
marital_status
DIVORCED
MARRIED
SEPARATED
SINGLE
WIDOWED
NA

How do demographics influence in-hospital mortality?

```
glm(formula = hospital_expire_flag ~ age + ethnicity + marital_status +
    language, family = "binomial", data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.1146  -0.4583  -0.3812  -0.3054   2.8384

Coefficients:
                                Estimate Std. Error z value Pr(>|z|)
(Intercept)                    -3.107213   0.651502  -4.769 1.85e-06 ***
age                             0.031763   0.001774  17.901  < 2e-16 ***
ethnicityHISPANIC OR LATINO    -0.013091   0.196425  -0.067 0.946863
ethnicityOTHER                 -0.016074   0.186942  -0.086 0.931477
ethnicityUNABLE TO OBTAIN       0.803709   0.151518   5.304 1.13e-07 ***
ethnicityUNKNOWN/NOT SPECIFIED  0.562160   0.159312   3.529 0.000418 ***
ethnicityWHITE                  0.041665   0.079084   0.527 0.598298
marital_statusMARRIED          -0.009904   0.088537  -0.112 0.910929
marital_statusSEPARATED         0.224446   0.213855   1.050 0.293935
marital_statusSINGLE            0.009709   0.094831   0.102 0.918449
marital_statusWIDOWED          -0.113735   0.102765  -1.107 0.268403
languageENGL                   -1.487467   0.630198  -2.360 0.018259 *
languagePTUN                   -0.754769   0.640661  -1.178 0.238753
languageRUSS                   -1.210058   0.642498  -1.883 0.059651 .
languageSPAN                   -1.311704   0.657075  -1.996 0.045904 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**' 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 15330  on 27223  degrees of freedom
Residual deviance: 14792  on 27209  degrees of freedom
  (17028 observations deleted due to missingness)
AIC: 14822
```

Standards

- “The wonderful thing about standards is that there are so many to choose from!”
- For example, consider prescriptions in MIMIC



https://xkcd.com/927/

Two Prescription

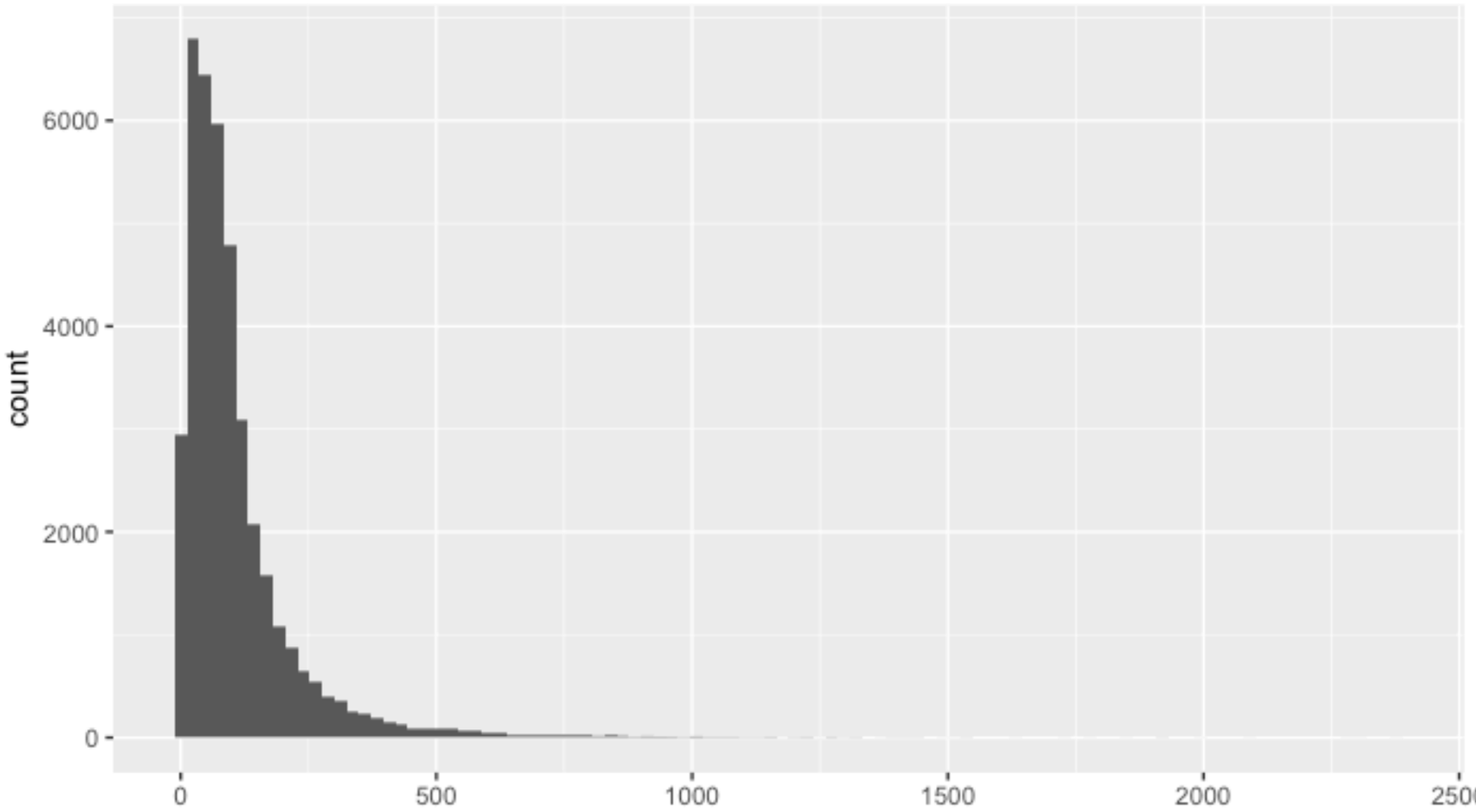
SUBJECT_ID	57139	57139
HADM_ID	155470	155470
ICUSTAY_ID	NA	NA
STARTDATE	2185-12-07	2185-12-07
ENDDATE	2185-12-07	2185-12-23
DRUG_TYPE	MAIN	MAIN
DRUG	Acetaminophen	Clobetasol Propionate 0.05%Cream
DRUG_NAME_POE	Acetaminophen	Clobetasol Propionate 0.05%Cream
DRUG_NAME_GENERIC	Acetaminophen	Clobetasol Propionate 0.05%Cream
FORMULARY_DRUG_CD	ACET325	CLOB.05C30
GSN	4489	7634
NDC	182844789	472040030
PROD_STRENGTH	325mg Tablet	30gm Tube
DOSE_VAL_RX	325-650	1
DOSE_UNIT_RX	mg	Appl
FORM_VAL_DISP	1-2	0.01
FORM_UNIT_DISP	TAB	TUBE
ROUTE	PO	TP

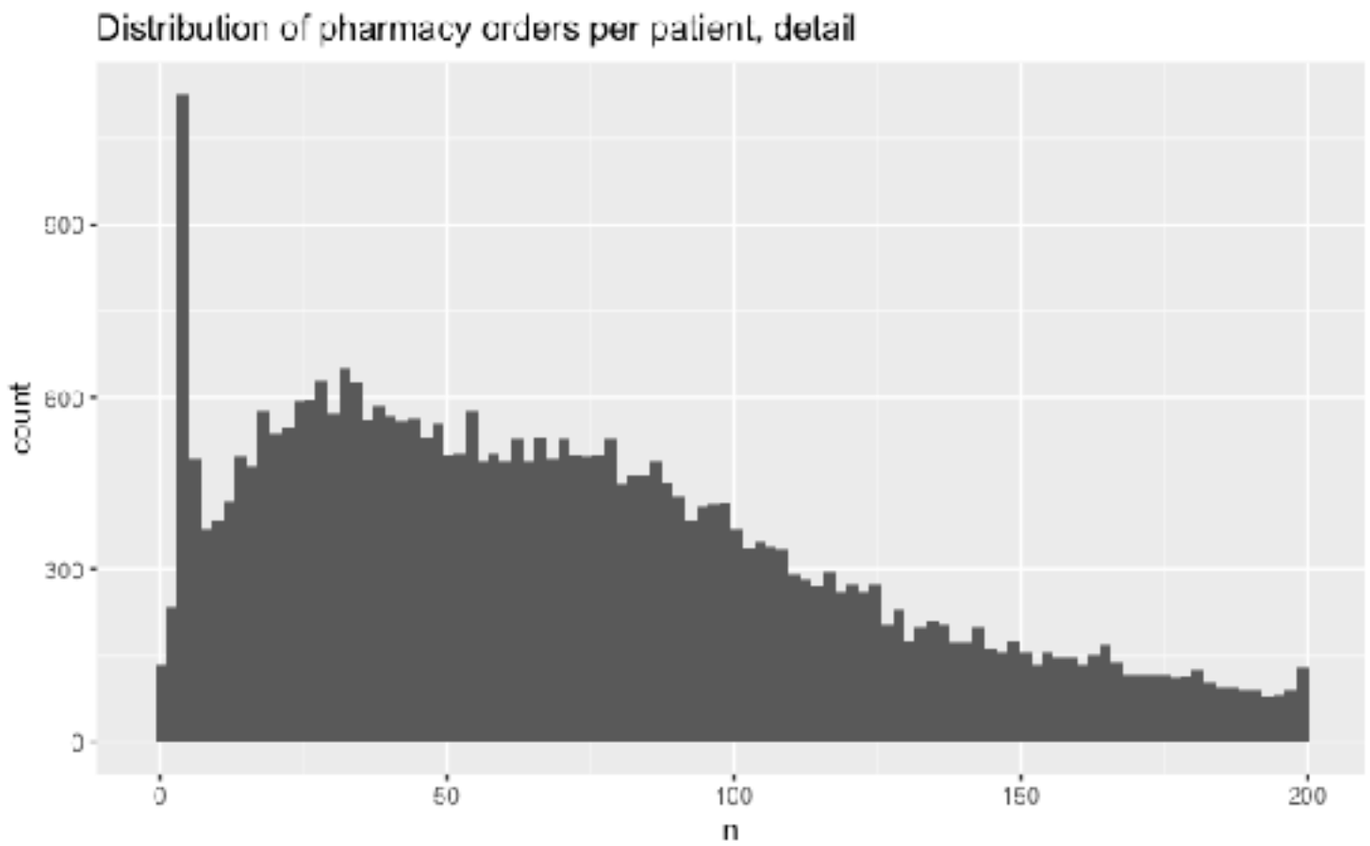
Most Common Prescriptions

	NDC Code	count
Iso-Osmotic Dextrose	0	86935
Sodium Chloride 0.9% Flush	0	83392
Insulin	0	81356
SW	0	72458
Magnesium Sulfate	409672924	55211
D5W	0	54938
Furosemide	517570425	53073
Potassium Chloride	338070341	47968
D5W	338001702	43038
LR	338011704	35407
Vancomycin	338355248	34741
0.9% Sodium Chloride	338004904	34682
Potassium Chloride	456066270	32533
Heparin	63323026201	31413
NS	338004902	30815

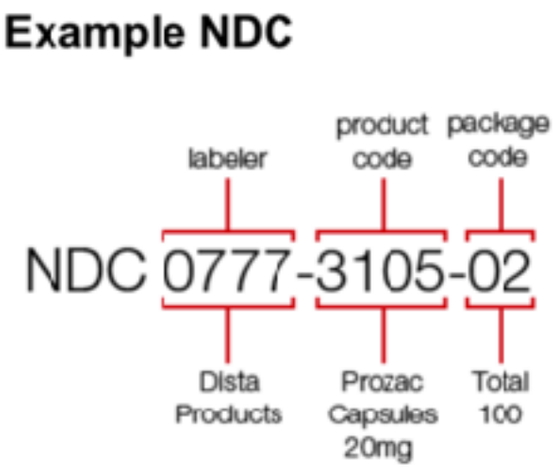
Next Most Common Prescriptions

	NDC Code	count
NS	338004903	29079
0.9% Sodium Chloride	338004903	28872
Metoprolol Tartrate	51079025520	28781
Insulin	88222033	26431
Pantoprazole	8084199	26379
Bag	0	25745
NS	338004904	25495
Vial	0	24497
Magnesium Sulfate	517260225	24212
5% Dextrose	338001702	24072
Potassium Chloride	58177020211	23881
Furosemide	74610204	23354
NS	338004938	23288
Potassium Chloride	58177000111	22976
Acetaminophen	182844789	22867


Distribution of pharmacy orders per admission
count
6000
4000
2000
0
0
500
1000
1500
2000
2500
n


Distribution of pharmacy orders per patient, detail
count
600
300
0
0
50
100
150
200
n

Medications



- NDC
  - "The Drug Listing Act of 1972 requires registered drug establishments to provide the Food and Drug Administration (FDA) with a current list of all drugs manufactured, prepared, propagated, compounded, or processed by it for commercial distribution. … Drug products are identified and reported using a unique, three-segment number, called the National Drug Code (NDC), which serves as a universal product identifier for drugs. FDA publishes the listed NDC numbers and the information submitted as part of the listing information in the NDC Directory which is updated daily.
- MedDRA
  - "the late 1990s, the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) developed MedDRA, a rich and highly specific standardised medical terminology to facilitate sharing of regulatory information internationally for medical products used by humans."
- RxNorm
  - "provides normalized names for clinical drugs and links its names to many of the drug vocabularies commonly used in pharmacy management and drug interaction software" and "and a tool for supporting semantic interoperation between drug terminologies and pharmacy knowledge base systems" (NLM)

Medications (more coding systems)

- Medicine Services and Procedures CPT Code range 90281- 99607
  - CPT Code range (90281-99607) for medicine contains CPT codes for immune globulins, serum or recombinant prods, immunization administration for vaccines/toxoids, vaccines, toxoids, psychiatry, biofeedback, dialysis, gastroenterology, ophthalmology, special otorhinolaryngologic services, cardiovascular, noninvasive vascular diagnostic studies, pulmonary, allergy and clinical immunology, endocrinology, neurology and neuromuscular procedures, central nervous system assessments/tests (neuro-cognitive, mental status, speech testing), health and behavior assessment/intervention, hydration, therapeutic, prophylactic, diagnostic injections and infusions, and chemotherapy and other highly complex drug or highly complex biologic agent administration, photodynamic therapy, special dermatological procedures, physical medicine and rehabilitation, medical nutrition therapy, acupuncture, osteopathic manipulative treatment, chiropractic manipulative treatment, education and training for patient self-management, non-face-to-face nonphysician services, special services, procedures and reports, other services and procedures, home health procedures/services, medication therapy management services.

Medications (more coding systems)

- 2019 Healthcare Common Procedure Coding System
  - HCPCS codes are used for billing Medicare & Medicaid patients
  - HCPCS J-Codes: Drugs administered other than oral method, chemotherapy drugs
  - These codes are used to report injectable drugs that ordinarily cannot be self-administered; chemotherapy, immunosuppressive drugs and inhalation solutions as well as some orally administered drugs.
- Commercial Coding Systems
  - The Generic Product Identifier (GPI) from Medi-Span is 14 characters made up of 7 couplets.
  - FDB [First Data Bank] has the Generic Sequence Number (GSN) also known as the Clinical Formulation ID or formerly as GCN Sequence Number (GCN Seq No), which is 6 digits in length. FDB also has the GCN (Formulation ID) which is 5 digits, …

What procedures were performed on the patient?

- PROCEDURES_ICD (n=240095)
- CPTEVENTS (n=573146)
- PROCEDUREEVENTS_MV (n=258066)

Most Common ICD9 Procedure Codes

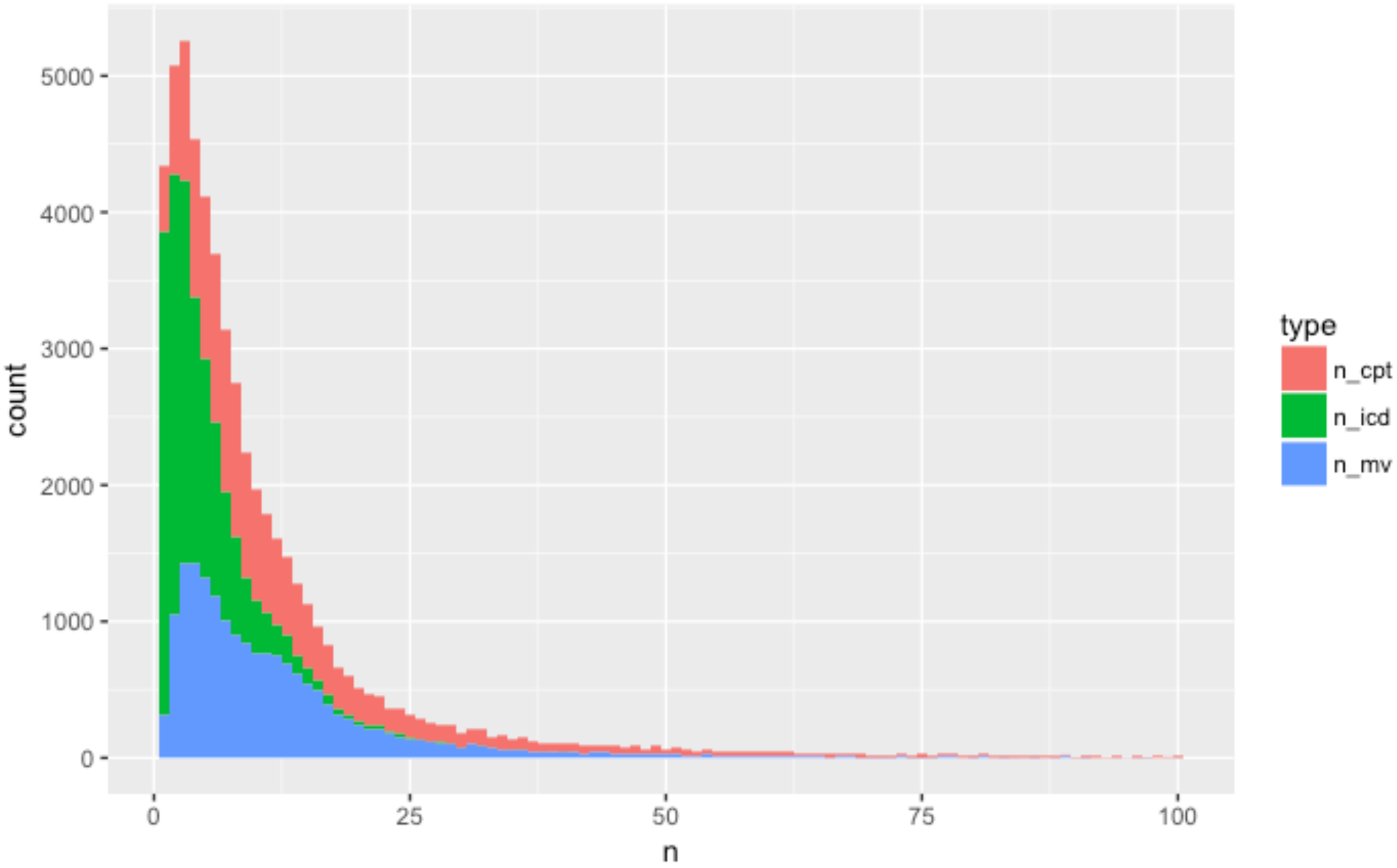
ICD9_code	n	Name
3893	14731	Venous catheterization, not elsewhere classified
9604	10333	Insertion of endotracheal tube
966	9300	Enteral infusion of concentrated nutritional substances
9671	9100	Continuous invasive mechanical ventilation for less than 96 consecutive hours
9904	7244	Transfusion of packed cells
3961	6838	Extracorporeal circulation auxiliary to open heart surgery
9672	6048	Continuous invasive mechanical ventilation for 96 consecutive hours or more
9955	5842	Prophylactic administration of vaccine against other diseases
8856	5337	Coronary arteriography using two catheters
3891	4737	Arterial catheterization
3615	4401	Single internal mammary-coronary artery bypass
9915	4244	Parenteral infusion of concentrated nutritional substances
8872	3548	Diagnostic ultrasound of heart
3722	3311	Left heart cardiac catheterization
3324	3269	Closed [endoscopic] biopsy of bronchus
3995	3254	Hemodialysis

Procedures (CPT)

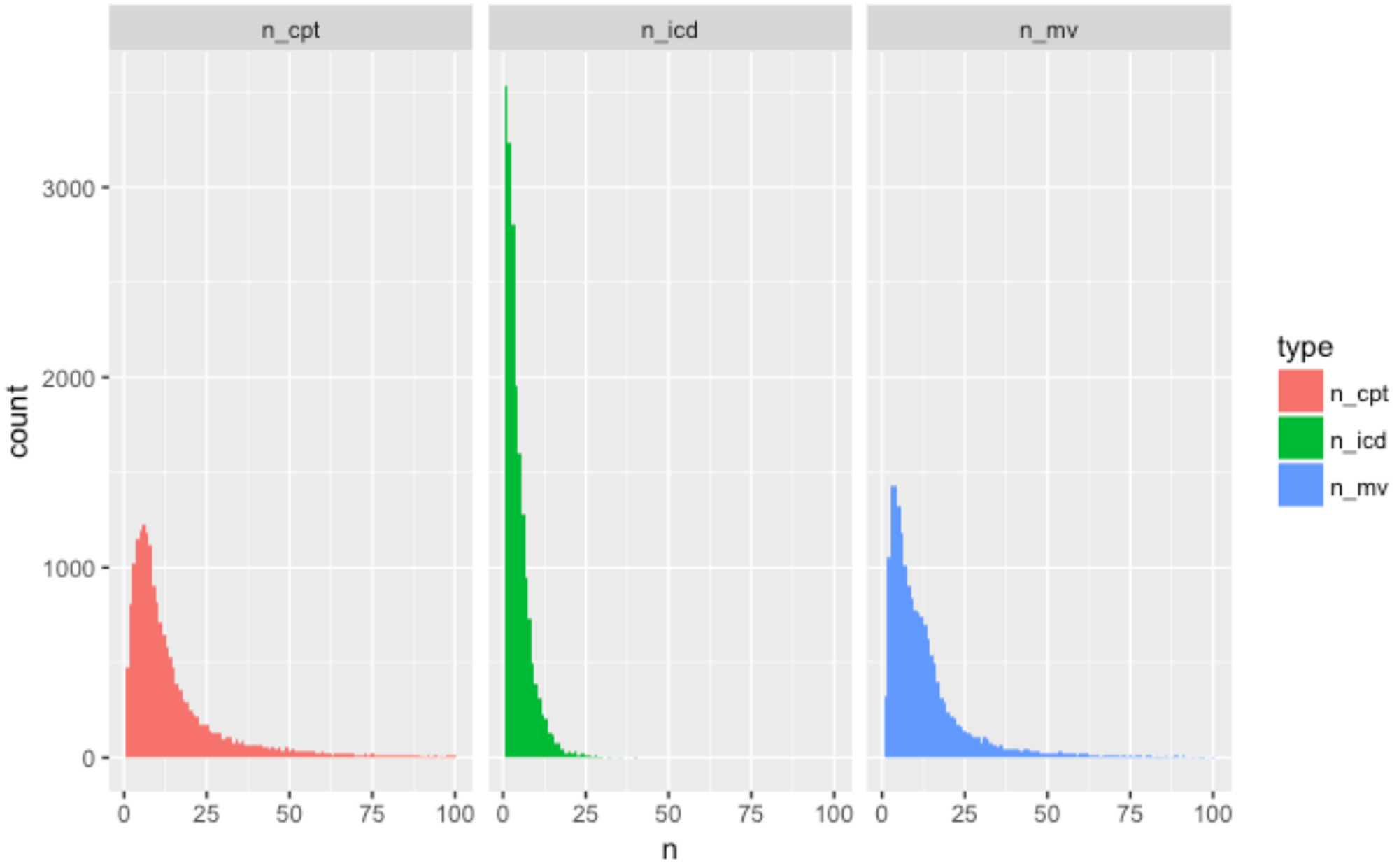
Medicine	90281–90399	Immune globulins, serum or recombinant prods
Medicine	90465–90474	Immunization administration for vaccines/toxoids
Medicine	90476–90749	Vaccines, toxoids
Medicine	90801–90899	Psychiatry
Medicine	90901–90911	Biofeedback
Medicine	90918–90925	End-Stage Renal Disease Services (deleted codes)
Medicine	90935–90999	Dialysis
Medicine	91000–91299	Gastroenterology
Medicine	92002–92499	Ophthalmology
Medicine	92502–92700	Special otorhinolaryngologic services
Medicine	92950–93799	Cardiovascular
Medicine	93875–93990	Noninvasive vascular diagnostic studies
Medicine	94002–94799	Pulmonary
Medicine	95004–95199	Allergy and clinical immunology
Medicine	95250–95251	Endocrinology
Medicine	95803–96020	Neurology and neuromuscular procedures
Medicine	96101–96125	Central nervous system assessments/tests (neuro-cogn:
Medicine	96150–96155	Health and behavior assessment/intervention
Medicine	96360–96549	Hydration, therapeutic, prophylactic, diagnostic inj
Medicine	96567–96571	Photodynamic therapy
Medicine	96900–96999	Special dermatological procedures
Medicine	97001–97799	Physical medicine and rehabilitation
Medicine	97802–97804	Medical nutrition therapy
Medicine	97810–97814	Acupuncture
Medicine	98925–98929	Osteopathic manipulative treatment
Medicine	98940–98943	Chiropractic manipulative treatment
Medicine	98960–98962	Education and training for patient self-management
Medicine	98966–98969	Non-face-to-face nonphysician services
Medicine	99000–99091	Special services, procedures and reports
Medicine	99170–99199	Other services and procedures
Medicine	99500–99602	Home health procedures/services
Medicine	99605–99607	Medication therapy management services

Surgery	10000–10022	General
Surgery	10040–19499	Integumentary system
Surgery	20000–29999	Musculoskeletal system
Surgery	30000–32999	Respiratory system
Surgery	33010–37799	Cardiovascular system
Surgery	38100–38999	Hemic and lymphatic systems
Surgery	39000–39599	Mediastinum and diaphragm
Surgery	40490–49999	Digestive system
Surgery	50010–53899	Urinary system
Surgery	54000–55899	Male genital system
Surgery	55920–55980	Reproductive system and intersex
Surgery	56340–56340	Laparoscopy, Surgical; Cholecystectomy
Surgery	56405–58999	Female genital system
Surgery	59000–59899	Maternity care and delivery
Surgery	60000–60699	Endocrine system
Surgery	61000–64999	Nervous system
Surgery	65091–68899	Eye and ocular adnexa
Surgery	69000–69979	Auditory system
Surgery	69990–69990	Operating microscope (deleted code)
Radiology	70000–76499	Diagnostic imaging
Radiology	76506–76999	Diagnostic ultrasound
Radiology	77001–77032	Radiologic guidance
Radiology	77051–77059	Breast mammography
Radiology	77071–77084	Bone/joint studies
Radiology	77261–77799	Radiation oncology
Radiology	78000–79999	Nuclear medicine

Procedure Codes per Admission



Procedure Codes per Admission



Lab measurements

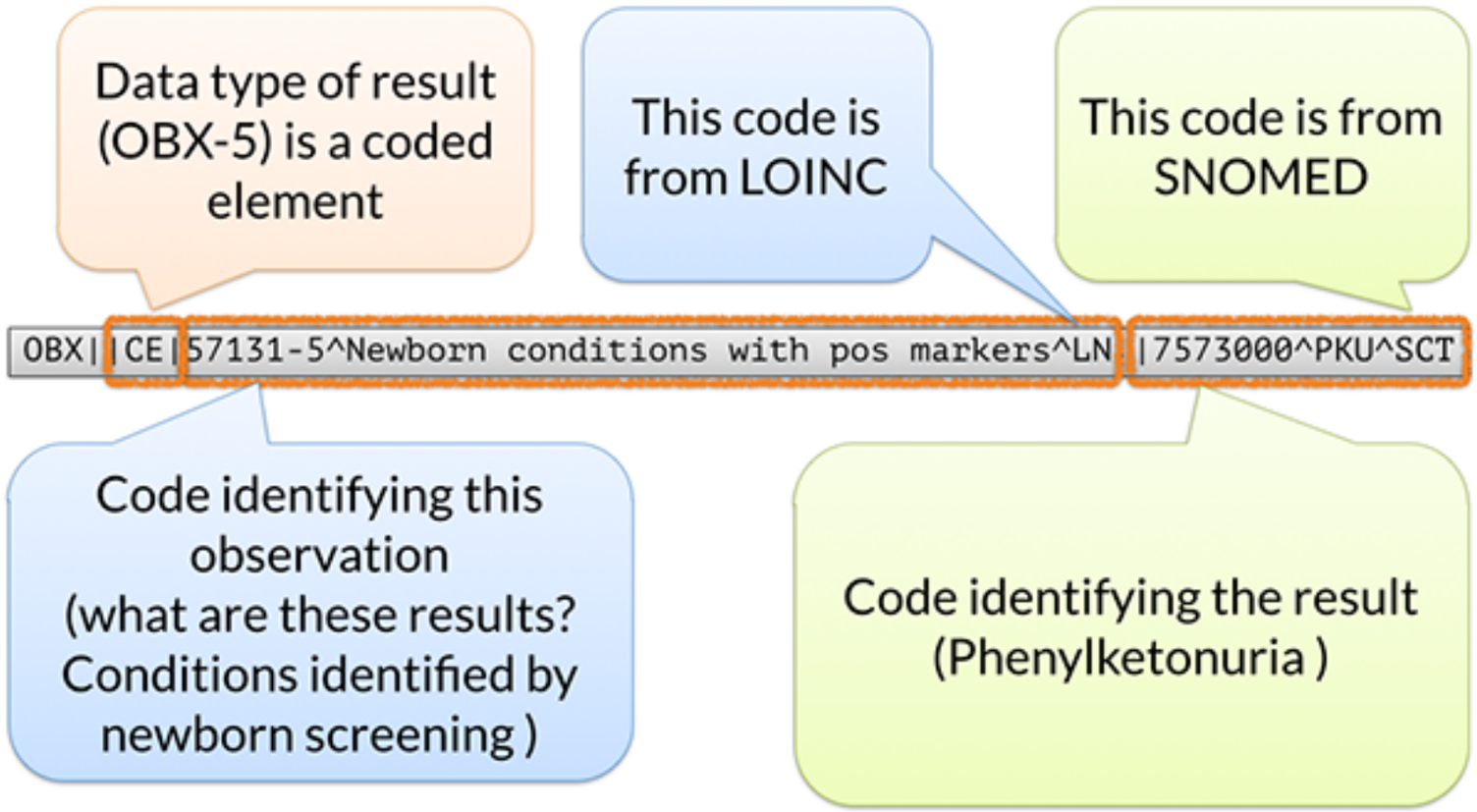
itemid	n	label	fluid	category	loinc
51221	881764	Hematocrit	Blood	Hematology	4544-3
50971	845737	Potassium	Blood	Chemistry	2823-3
50983	808401	Sodium	Blood	Chemistry	2951-2
50912	797389	Creatinine	Blood	Chemistry	2160-0
50902	795480	Chloride	Blood	Chemistry	2075-0
51006	791838	Urea Nitrogen	Blood	Chemistry	3094-0
50882	780648	Bicarbonate	Blood	Chemistry	1963-8
51265	778365	Platelet Count	Blood	Hematology	777-3
50868	769810	Anion Gap	Blood	Chemistry	1863-0
51301	753221	White Blood Cells	Blood	Hematology	804-5
51222	752444	Hemoglobin	Blood	Hematology	718-7
50931	748896	Glucose	Blood	Chemistry	2345-7
51249	748147	MCHC	Blood	Hematology	786-4
51279	747999	Red Blood Cells	Blood	Hematology	789-8
51248	747994	MCH	Blood	Hematology	785-6
51250	747977	MCV	Blood	Hematology	787-2

Labs for patient 2, admission 163353

subj	hadm	item	time	value	units	flag	label	fluid	categ	loinc
2	163353	51143	2138-07-17 20:48:00	0.00	%	NA	Atypical Lymphocytes	Blood	Hem	733-6
2	163353	51144	2138-07-17 20:48:00	0.00	%	NA	Bands	Blood	Hem	763-3
2	163353	51146	2138-07-17 20:48:00	0.00	%	NA	Basophils	Blood	Hem	704-7
2	163353	51200	2138-07-17 20:48:00	0.00	%	NA	Eosinophils	Blood	Hem	711-2
2	163353	51221	2138-07-17 20:48:00	0.00	%	abnormal	Hematocrit	Blood	Hem	4544-3
2	163353	51222	2138-07-17 20:48:00	0.00	g/dL	abnormal	Hemoglobin	Blood	Hem	718-7
2	163353	51244	2138-07-17 20:48:00	0.00	%	NA	Lymphocytes	Blood	Hem	731-0
2	163353	51248	2138-07-17 20:48:00	0.00	pg	abnormal	MCH	Blood	Hem	785-6
2	163353	51249	2138-07-17 20:48:00	0.00	%	abnormal	MCHC	Blood	Hem	786-4
2	163353	51250	2138-07-17 20:48:00	0.00	fL	abnormal	MCV	Blood	Hem	787-2
2	163353	51251	2138-07-17 20:48:00	0.00	%	NA	Metamyelocytes	Blood	Hem	28541-1
2	163353	51254	2138-07-17 20:48:00	0.00	%	NA	Monocytes	Blood	Hem	742-7
2	163353	51255	2138-07-17 20:48:00	0.00	%	NA	Myelocytes	Blood	Hem	26498-6
2	163353	51256	2138-07-17 20:48:00	100.00	%	NA	Neutrophils	Blood	Hem	761-7
2	163353	51265	2138-07-17 20:48:00	5.00	K/uL	abnormal	Platelet Count	Blood	Hem	777-3

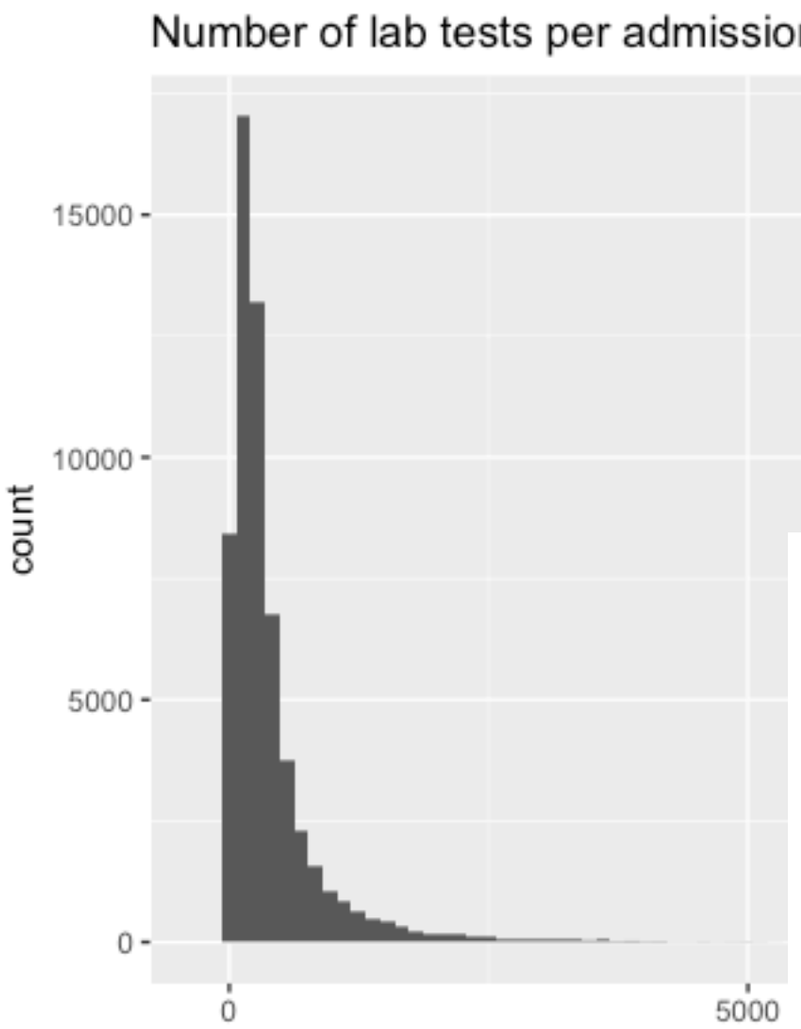
Reporting lab results
Logical Observation Identifiers Names and Codes

Most laboratory and clinical systems today are sending data out using the HL7 version 2 messaging standard. Looking at an example of the place in the HL7 message where the test results go, you can see how a LOINC code identifies the question and a SNOMED CT code represents the answer:



- Component (Analyte)
- Property
- Time
- System (Specimen)
- Scale
- Method

Lab tests per admission



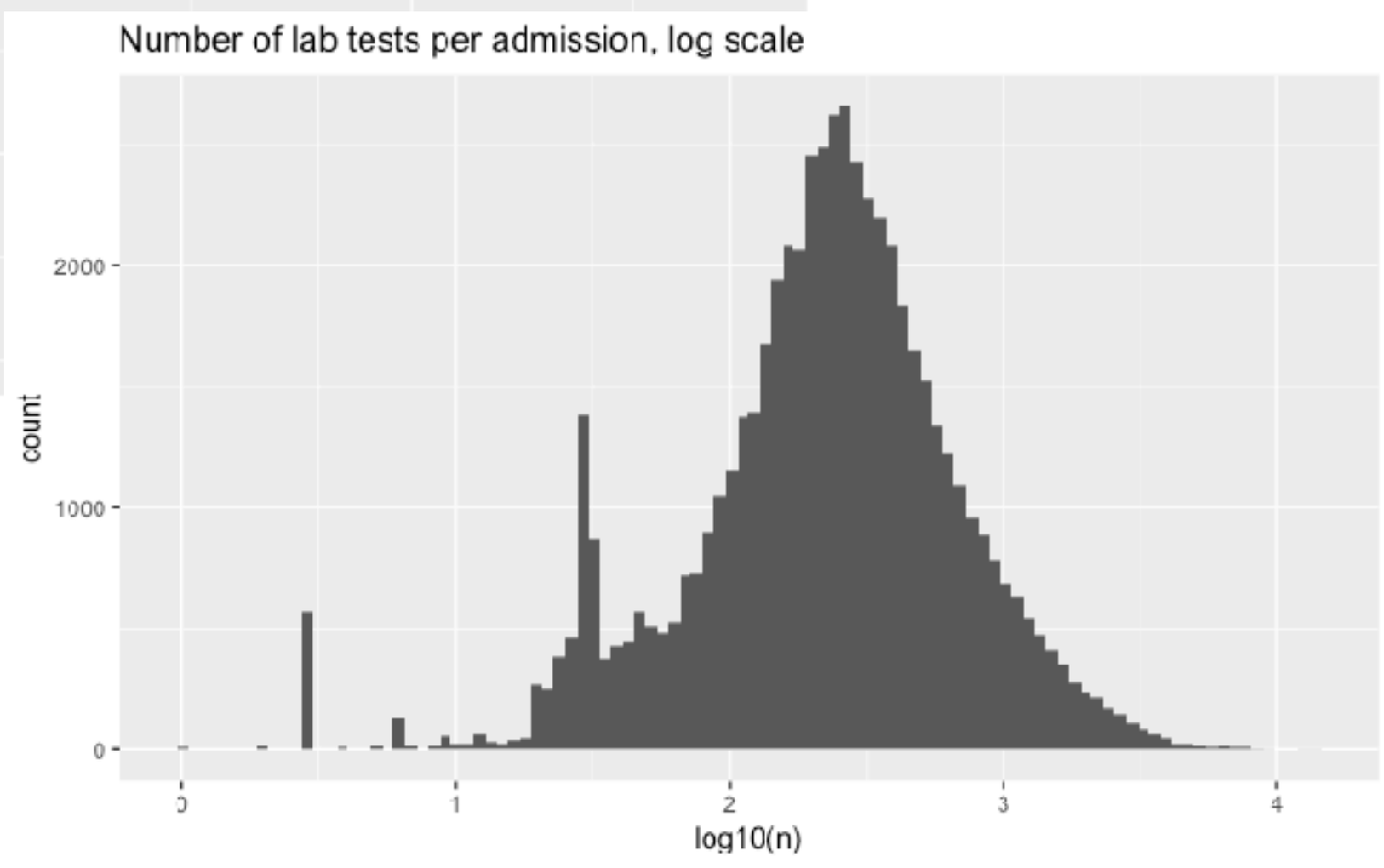
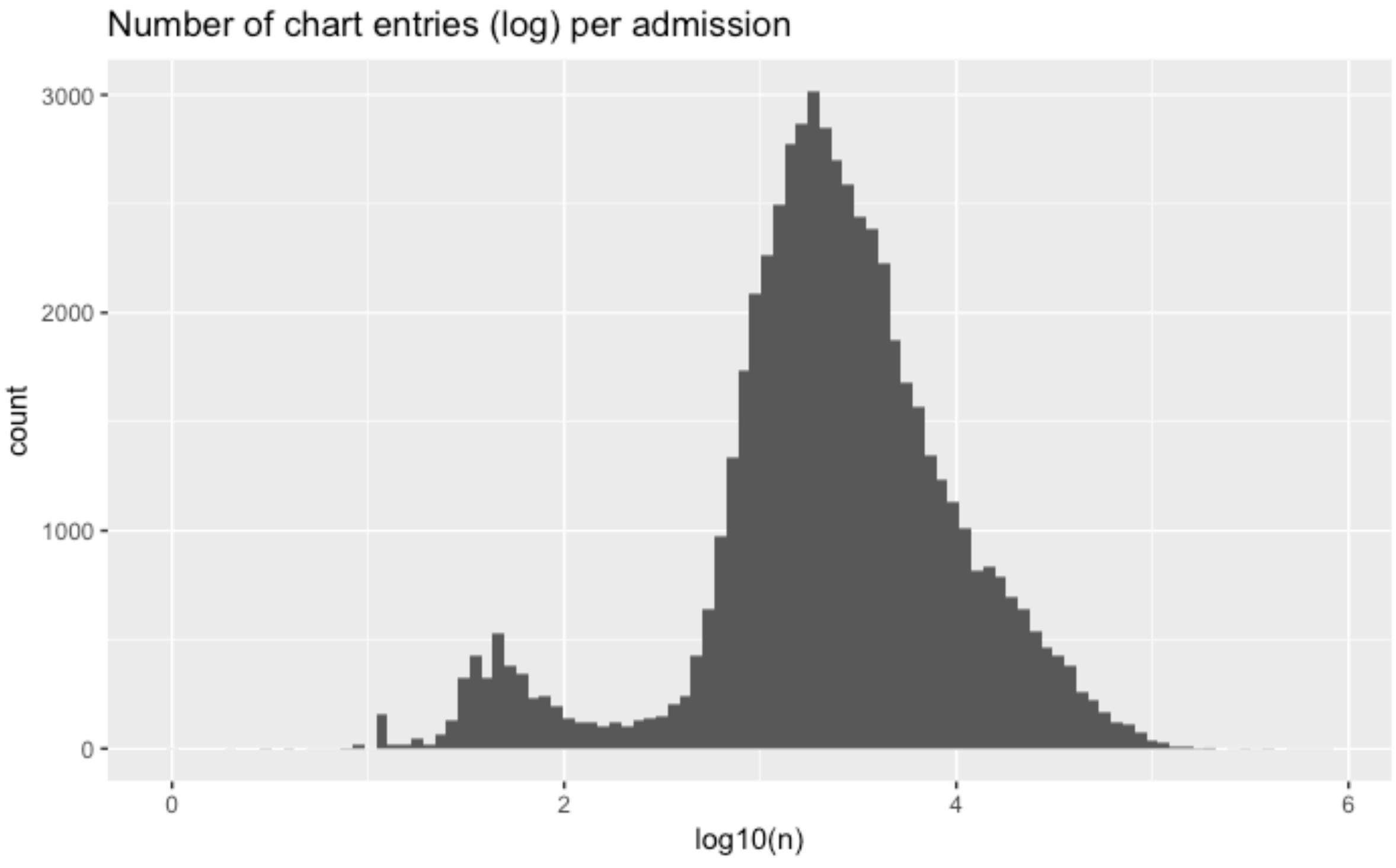


Chart Events

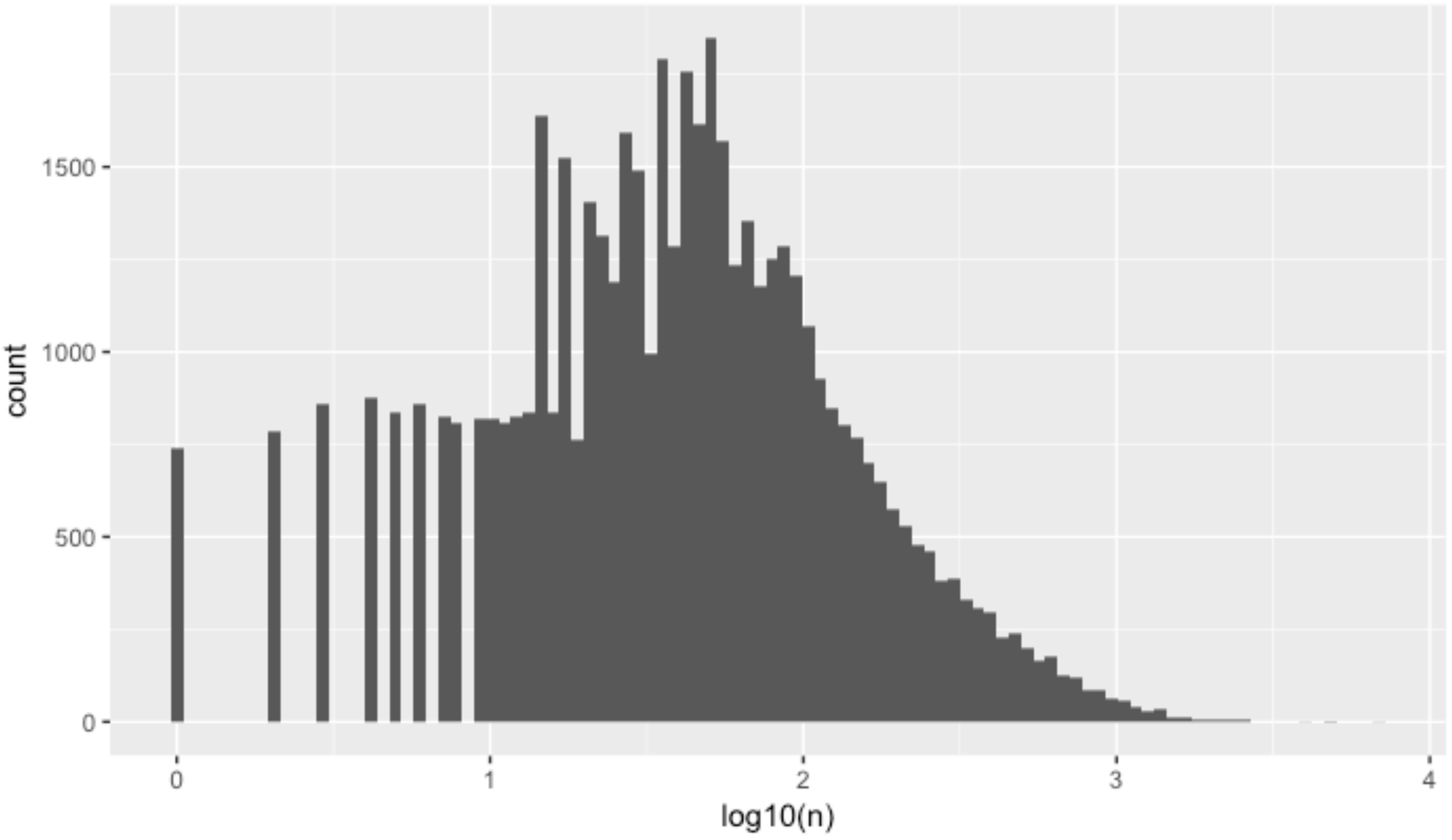
itemid	n	label	category	units	param_type
211	5180809	Heart Rate	NA	NA	NA
742	3464326	calprevflg	NA	NA	NA
646	3418917	SpO2	NA	NA	NA
618	3386719	Respiratory Rate	NA	NA	NA
212	3303151	Heart Rhythm	NA	NA	NA
161	3236350	Ectopy Type	NA	NA	NA
128	3216866	Code Status	NA	NA	NA
550	3205052	Precautions	NA	NA	NA
1125	2955851	Service Type	NA	NA	NA
220045	2762225	Heart Rate	Routine Vital Signs	bpm	Numeric
220210	2737105	Respiratory Rate	Respiratory	insp/min	Numeric
220277	2671816	O2 saturation pulseoxymetry	Respiratory	%	Numeric
159	2544519	Ectopy Frequency	NA	NA	NA
1484	2261065	Risk for Falls	NA	NA	NA
51	2096678	Arterial BP [Systolic]	NA	NA	NA
8368	2085994	Arterial BP [Diastolic]	NA	NA	NA

Chart entries



Outputs

itemid	n	label	category	units
40055	1917421	Urine Out Foley	NA	NA
226559	1186717	Foley	Output	mL
40076	152716	Chest Tubes CTICU CT 1	NA	NA
43175	108982	Urine .	NA	NA
40054	81828	Stool Out Stool	NA	NA
226588	81128	Chest Tube #1	Output	mL
40069	69467	Urine Out Void	NA	NA


Number of outputs (log) per admission
count
1500
1000
500
0
0
1
2
3
4
log10(n)

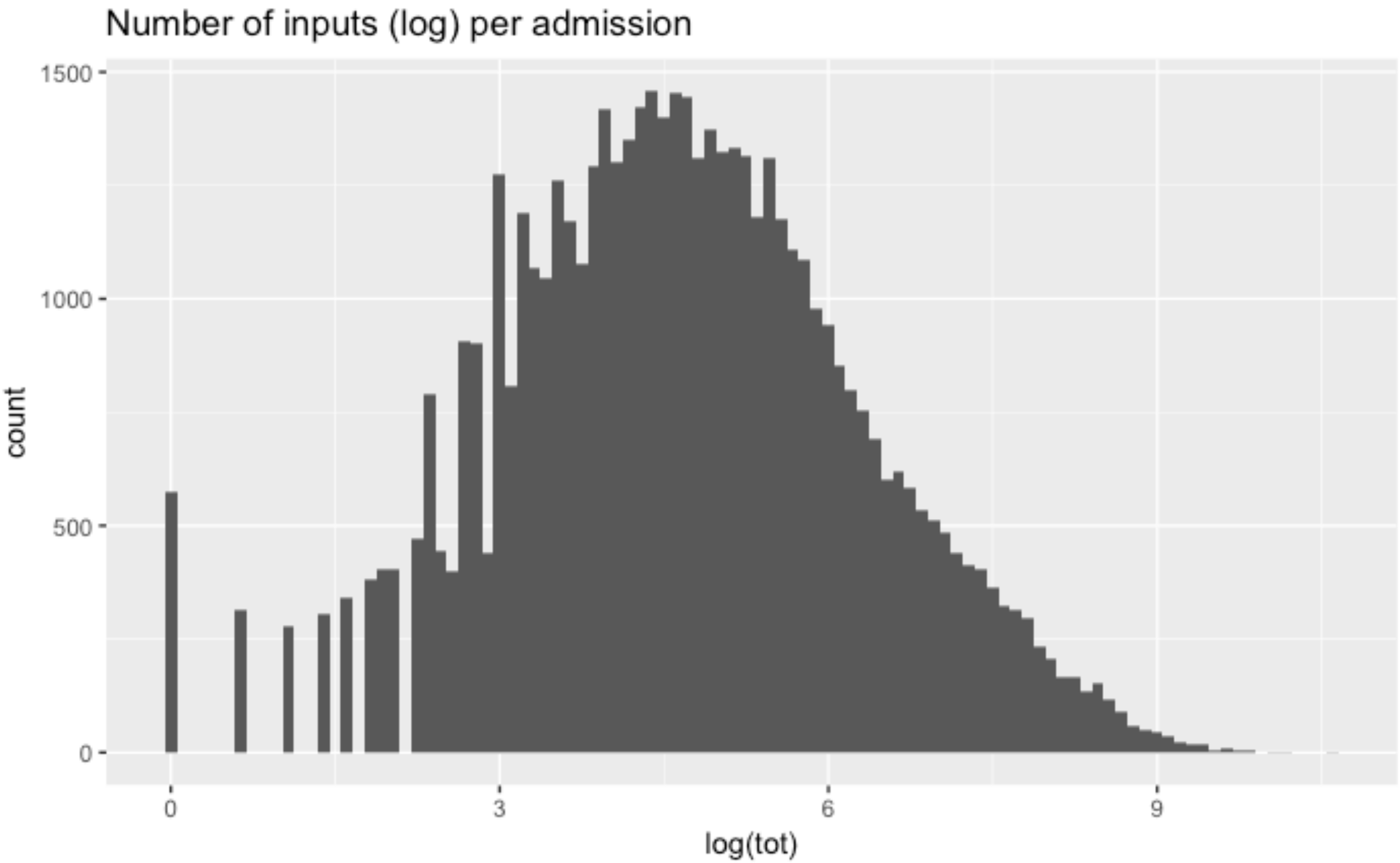
Inputs (CareVue)

itemid	n	label
30013	2557507	D5W
30018	2392372	.9% Normal Saline
30131	924614	Propofol
30045	825758	Insulin
30025	813242	Heparin
30118	780555	Fentanyl
30128	554582	Neosynephrine-k
30124	505509	Midazolam
30120	476971	Levophed-k
30140	373023	N/A

Inputs (MetaVision)

itemid	n	label	category	unit	param_type
225158	527855	NaCl 0.9%	Fluids/Intake	mL	Solution
220949	406345	Dextrose 5%	Fluids/Intake	mL	Solution
225943	246312	Solution	Fluids/Intake	mL	Solution
222168	178819	Propofol	Medications	mg	Solution
226452	135438	PO Intake	Fluids/Intake	mL	Solution
223258	119668	Insulin - Regular	Medications	units	Solution
225799	97629	Gastric Meds	Fluids/Intake	mL	Solution
221749	93571	Phenylephrine	Medications	mg	Solution
221906	89697	Norepinephrine	Medications	mg	Solution
221744	86340	Fentanyl	Medications	mg	Solution

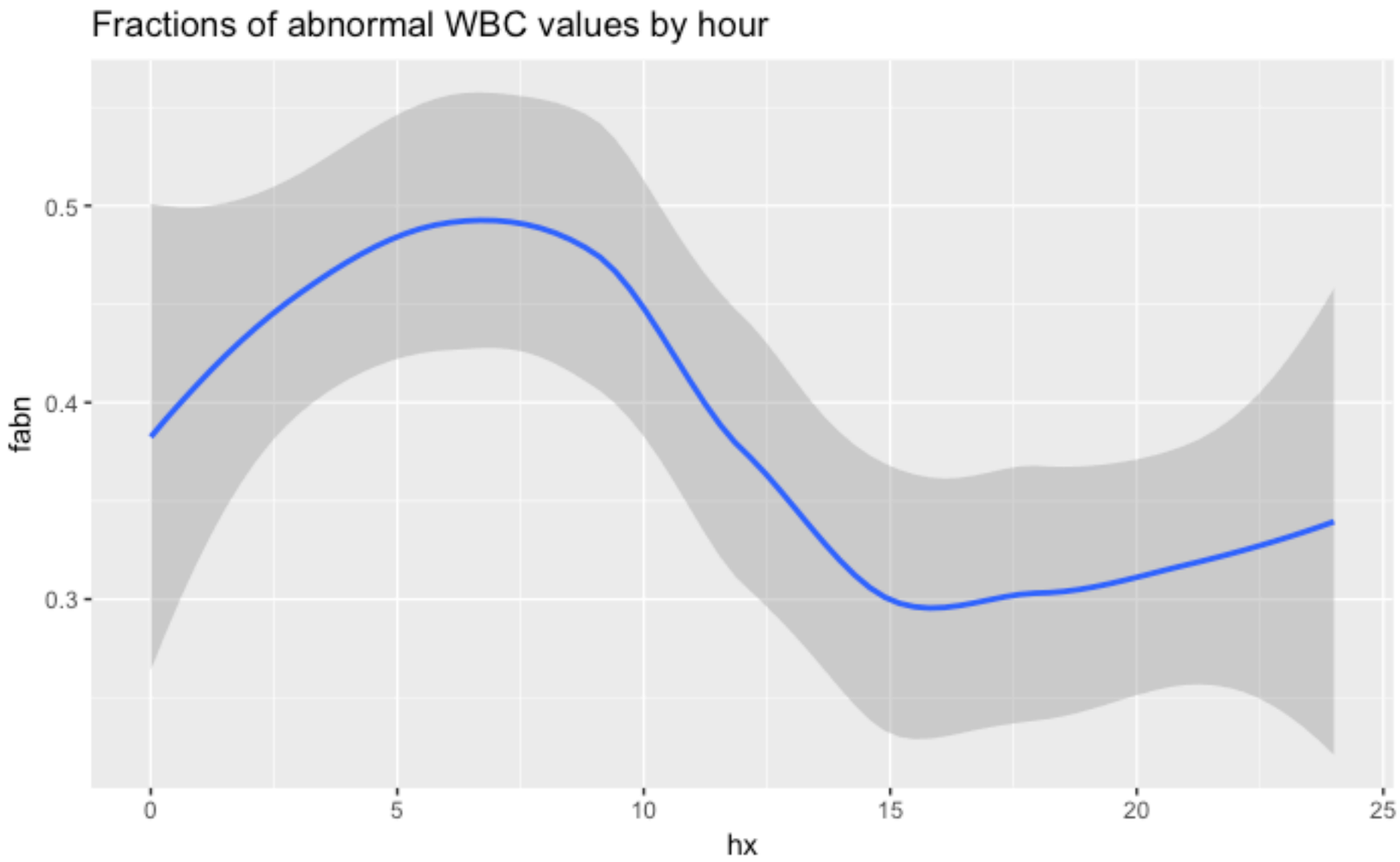
Inputs (combined MV and CV)



Biases in electronic health record data due to processes within the healthcare system: retrospective observational study

Denis Agniel,[1] Isaac S Kohane,[1,2] Griffin M Weber[1,3]

- Showed that for many lab results, “process measures” of the data are more important than actual values in predicting outcomes
  - E.g., White Blood Cell count

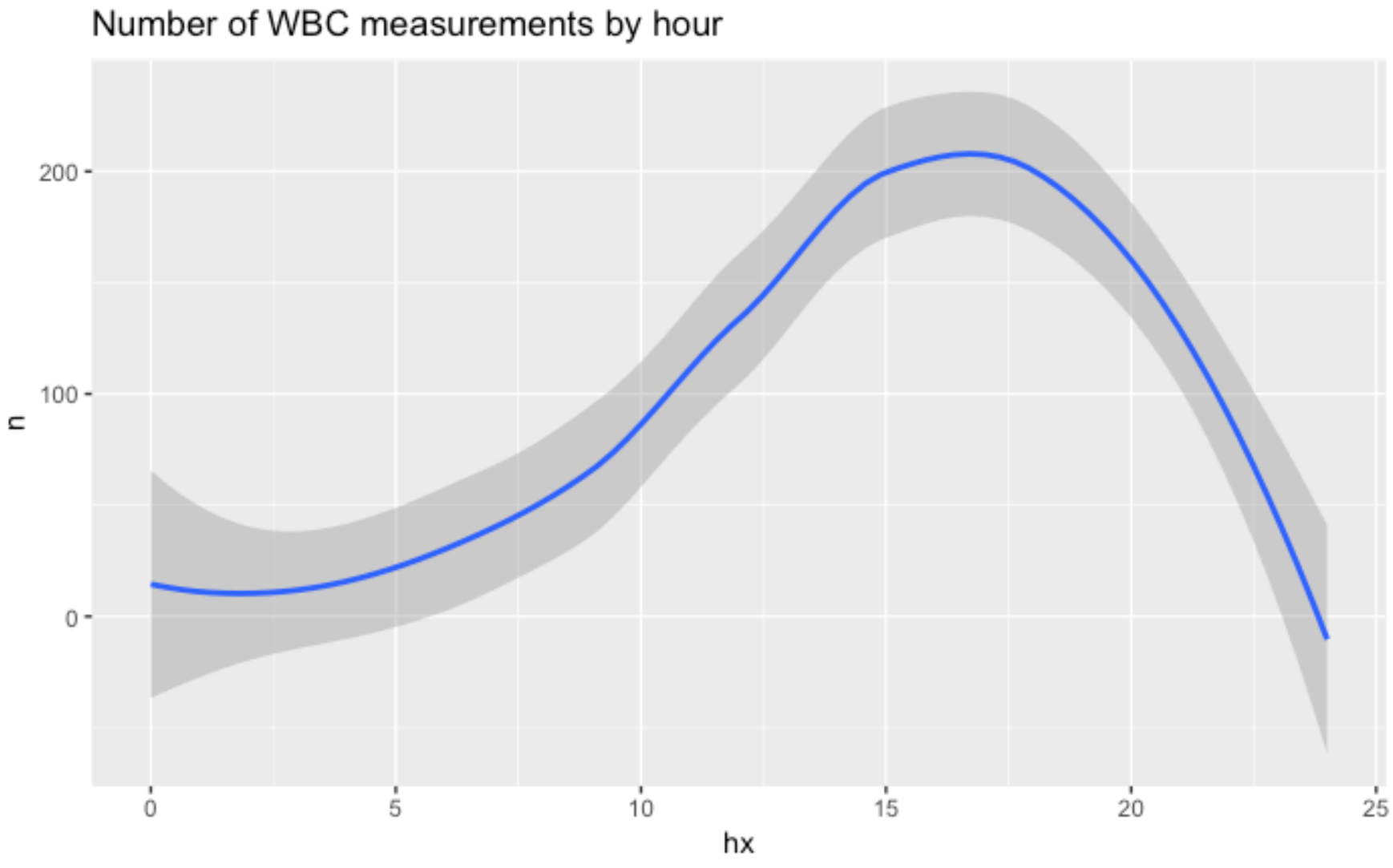


LR model to predict mortality from number of WBC measurements and number abnormal, per hour

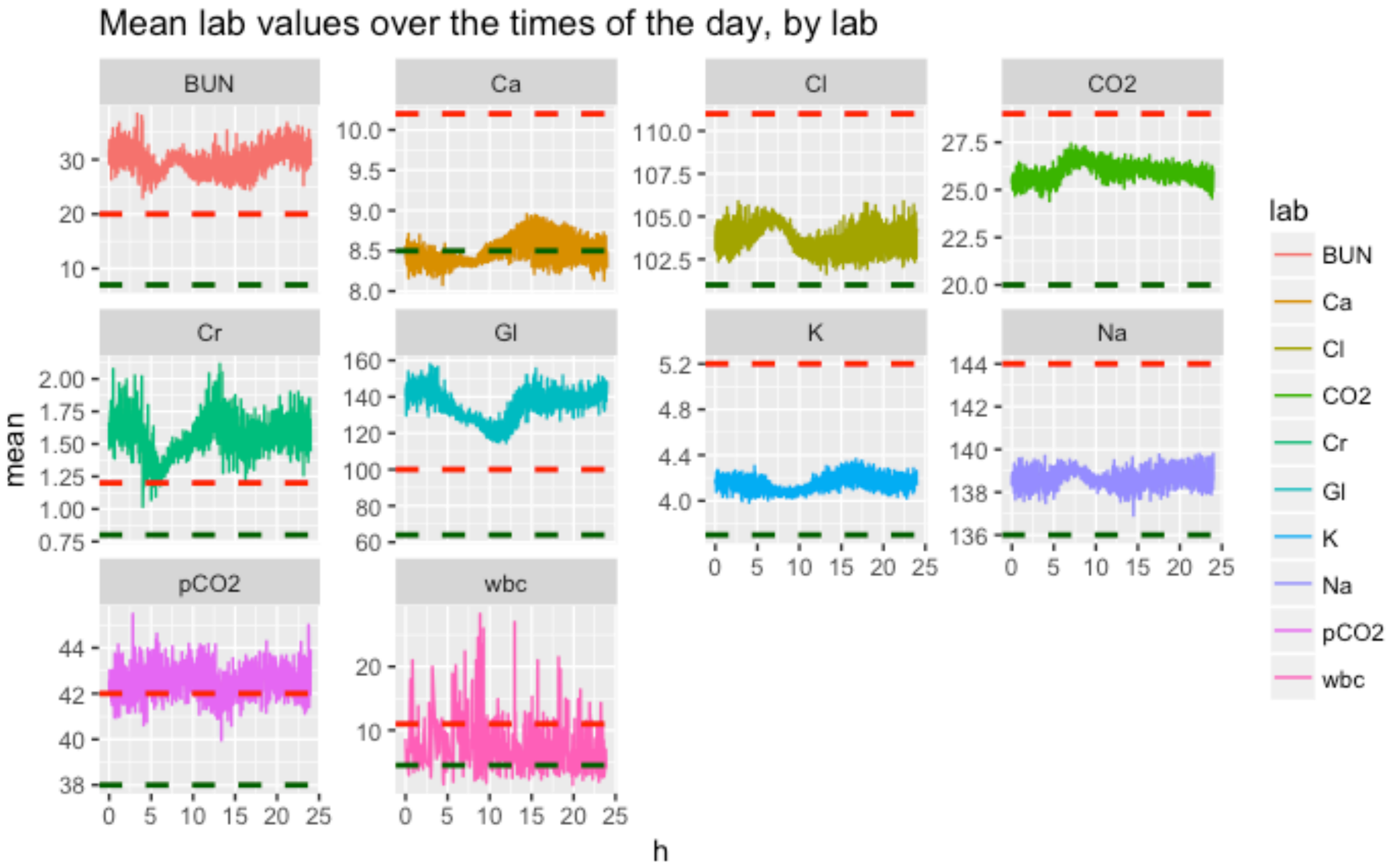
```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.8045  -1.0958  -0.5012   1.1245   2.3401

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.04321    0.11487   0.376 0.706758
H0            0.75871    0.88579   0.857 0.391700
H1            0.45657    0.76061   0.600 0.548333
H2            0.39502    0.65687   0.601 0.547597
H3           15.46281  413.03082   0.037 0.970136
H4            0.87956    0.90070   0.977 0.328804
H5            0.19184    0.92995   0.206 0.836562
H6            0.43533    0.65352   0.666 0.505330
H7            0.05389    0.40893   0.132 0.895147
H8            1.36632    0.47436   2.880 0.003972 **
H9            0.07131    0.24685   0.289 0.772685
H10           0.02999    0.16509   0.182 0.855845
H11          -1.03418    0.32225  -3.209 0.001331 **
H12           0.15791    0.21427   0.737 0.461133
H13          -0.39467    0.31470  -1.254 0.209803
H14          -0.19412    0.18526  -1.048 0.294726
H15          -0.42509    0.15821  -2.687 0.007212 **
H16           0.24009    0.12191   1.969 0.048900 *
H17          -0.10166    0.15254  -0.666 0.505139
H18          -0.10116    0.18002  -0.562 0.574149
H19          -0.23376    0.24193  -0.966 0.333919
H20          -0.12929    0.18466  -0.700 0.483827
H21          -0.79920    0.27154  -2.943 0.003248 **
H22          -0.56242    0.36065  -1.559 0.118893
H23          -0.45735    0.47557  -0.962 0.336199
H24           0.08659    0.71026   0.122 0.902962
HA0          -1.78217    1.32944  -1.341 0.180071
HA1          -0.80485    1.28716  -0.625 0.531782
HA2          -1.39389    1.36913  -1.018 0.308639
HA3         -15.69112  413.03210  -0.038 0.969696
HA4          -0.91247    1.21520  -0.751 0.452723
HA5          -0.32100    1.38380  -0.232 0.816564
HA6          -1.32274    1.04715  -1.263 0.206524
HA7          -0.71769    0.93684  -0.766 0.443632
HA8          -1.71813    0.66992  -2.565 0.010327 *
HA9          -0.67054    0.51100  -1.312 0.189450
HA10         -0.19831    0.45897  -0.432 0.665693
HA11          1.72924    0.52482   3.295 0.000984 ***
HA12          0.03971    0.59225   0.067 0.946540
HA13          0.94444    0.62952   1.500 0.133550
HA14          0.22134    0.45705   0.484 0.628188
HA15          1.25147    0.44487   2.813 0.004906 **
HA16          0.04059    0.39246   0.103 0.917633
HA17          0.18535    0.46846   0.396 0.692352
HA18          0.49504    0.44025   1.124 0.260823
HA19         -0.02478    0.45548  -0.054 0.956612
HA20          0.41568    0.53548   0.776 0.437594
HA21          1.60231    0.60935   2.630 0.008550 **
HA22          0.52832    0.56629   0.933 0.350848
HA23          0.92591    0.88156   1.050 0.293580
HA24          0.67132    1.68820   0.398 0.690887
---
Signif. codes:  0 â€˜***â€™ 0.001 â€˜**â€™ 0.01 â€˜*â€™ 0.05
```

Relationship of WBC measurements at night to mortality

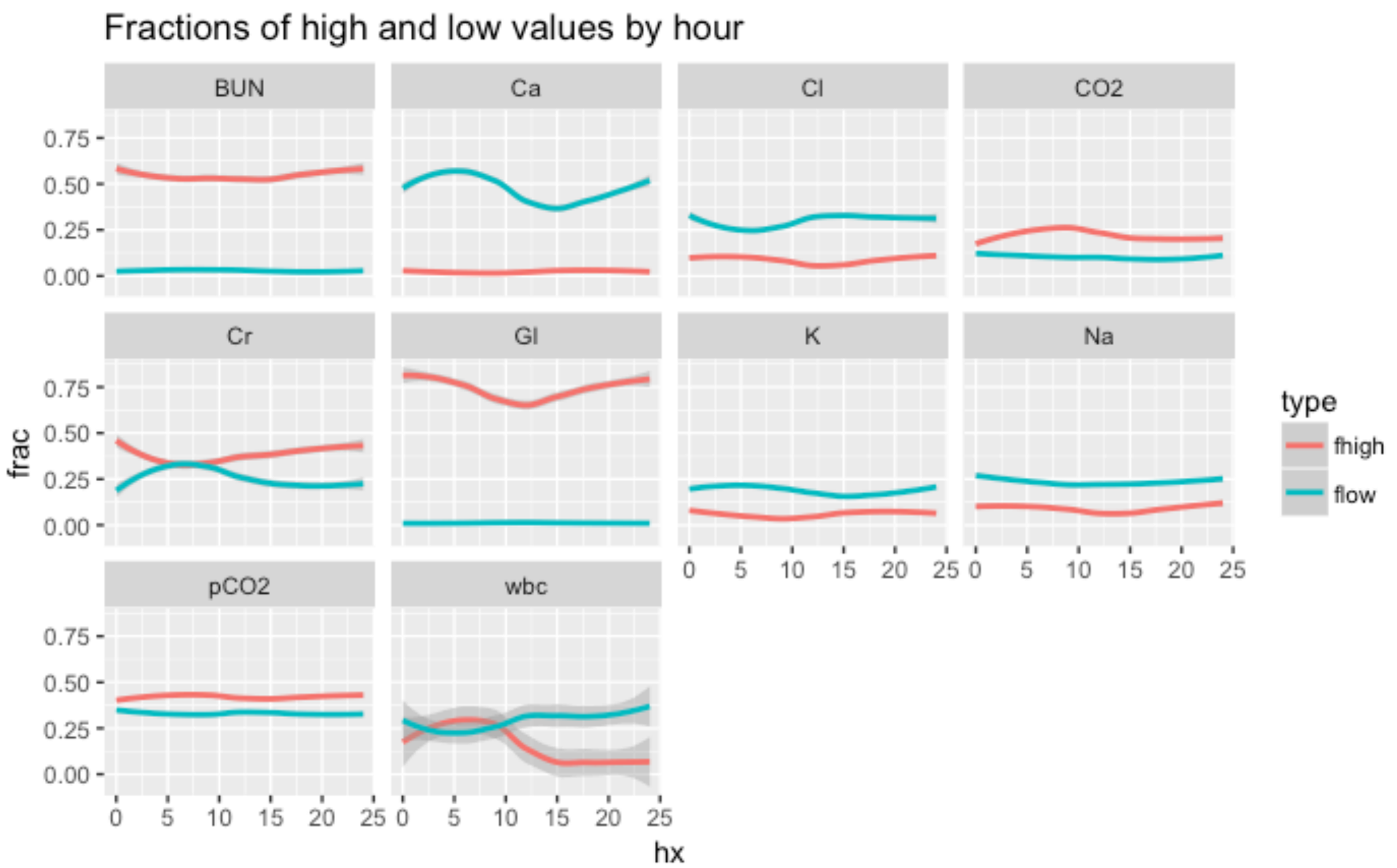
	0	1	2	3	4
FALSE	289	65	8	1	1
TRUE	253	67	3	0	0



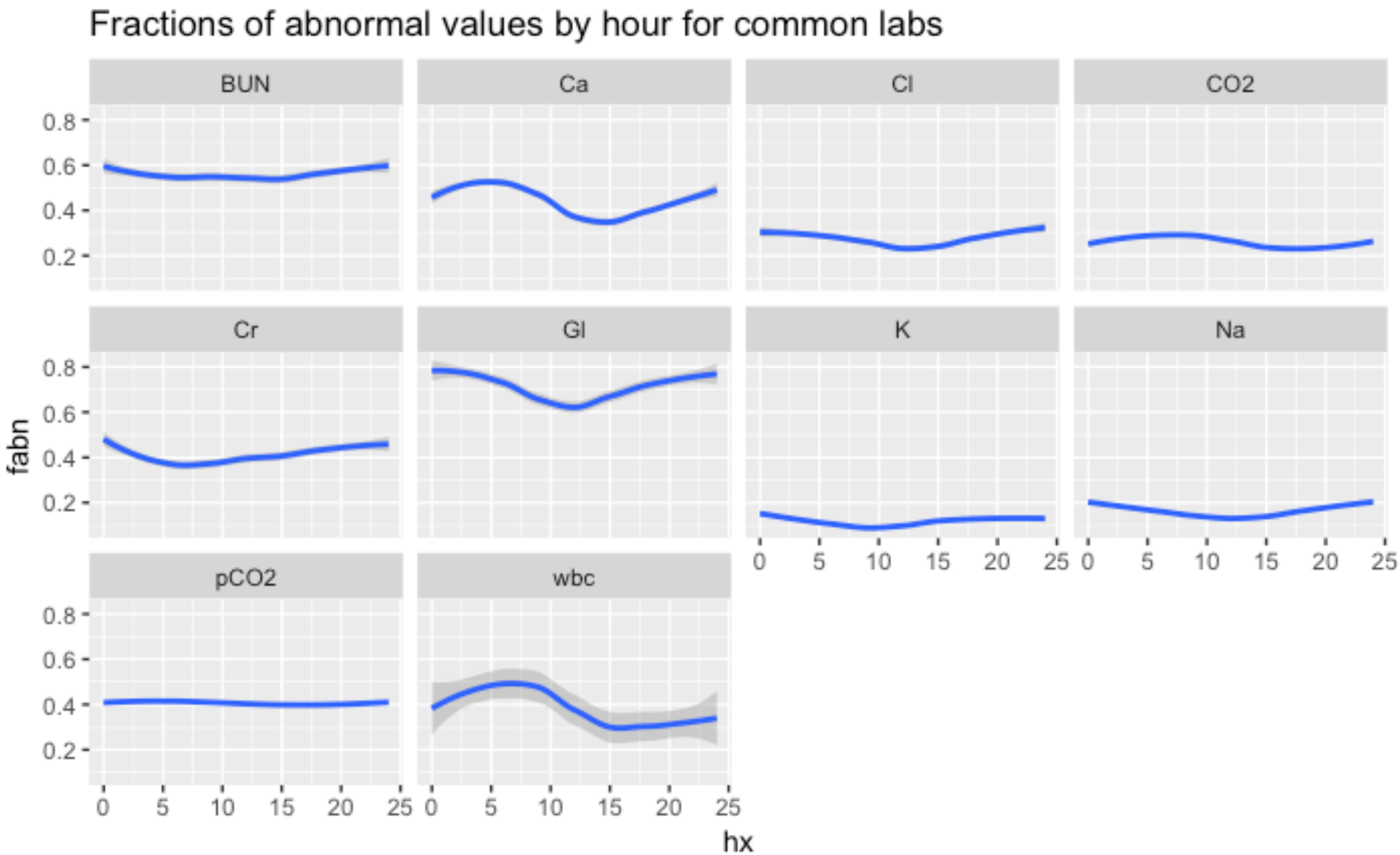
Lab values *do* vary by time of day

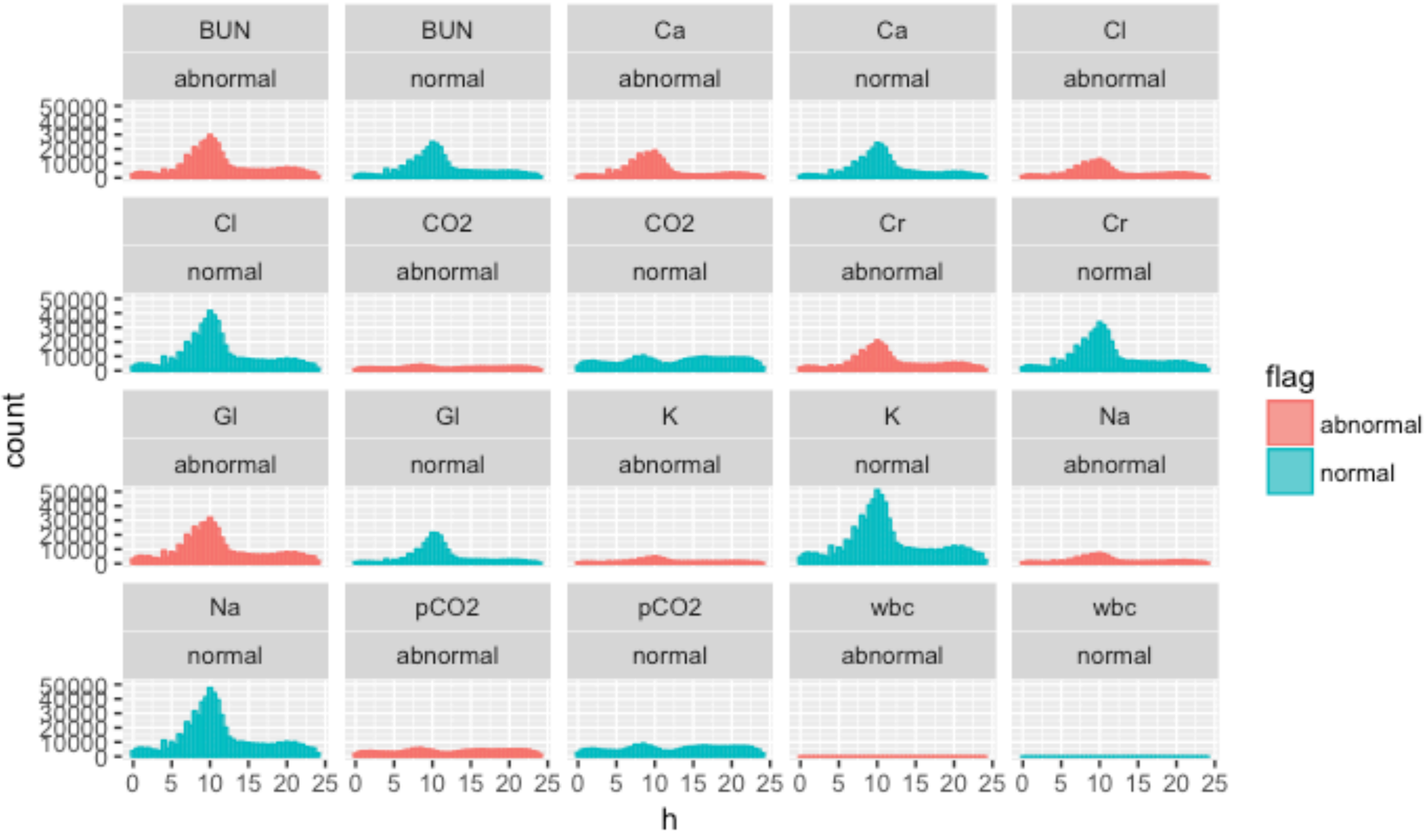


Fractions of high and low lab values *do* vary by hour

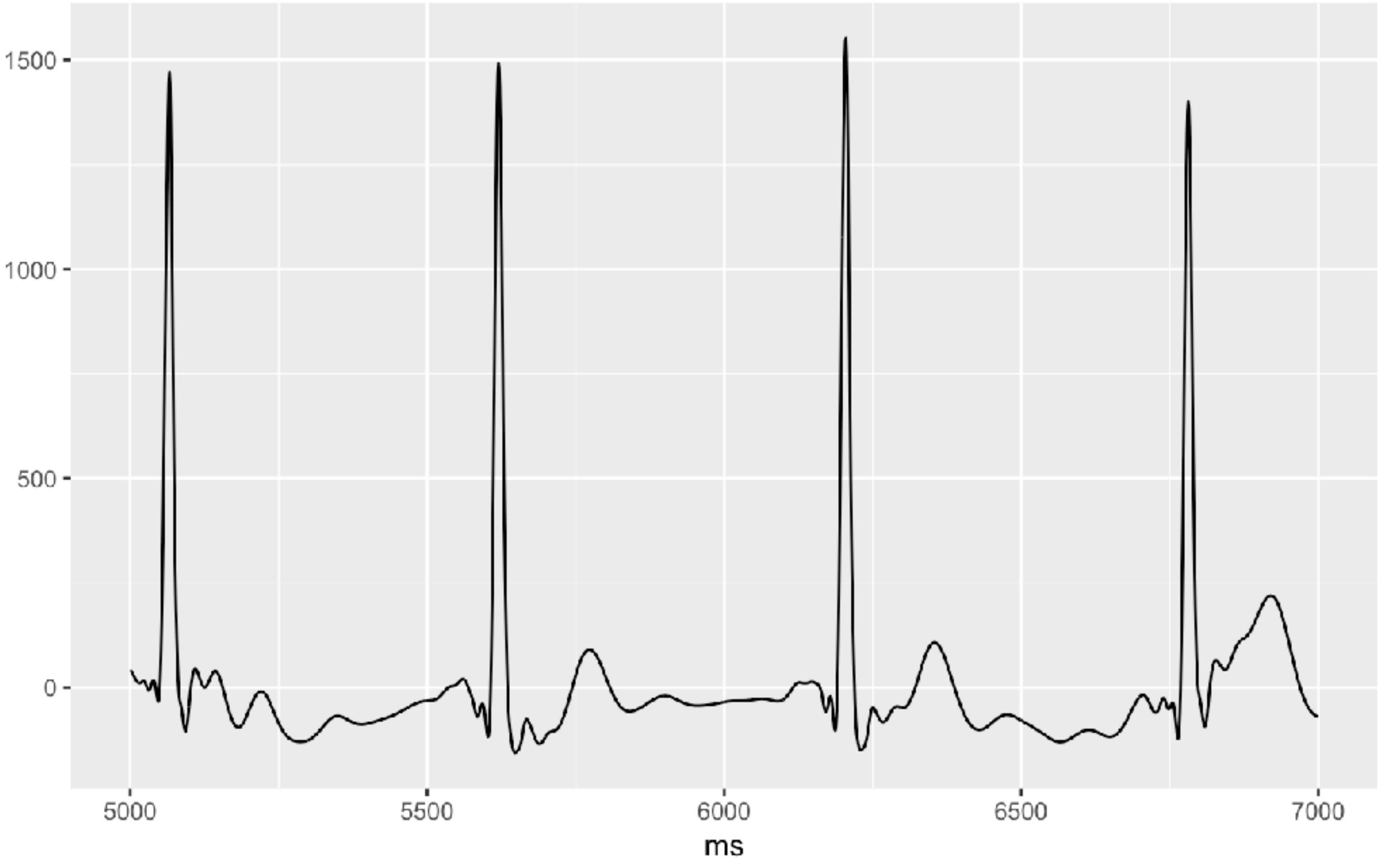


Fractions of abnormal lab values *do* vary by hour




Times of lab measurements, faceted by type
BUN
abnormal
BUN
normal
Ca
abnormal
Ca
normal
Cl
abnormal
Cl
normal
CO2
abnormal
CO2
normal
Cr
abnormal
Cr
normal
Gl
abnormal
Gl
normal
K
abnormal
K
normal
Na
abnormal
Na
normal
pCO2
abnormal
pCO2
normal
wbc
abnormal
wbc
normal
50000
40000
30000
20000
10000
0
0 5 10 15 20 25
h
flag
abnormal
normal

Data from wearables



Heart rate variability

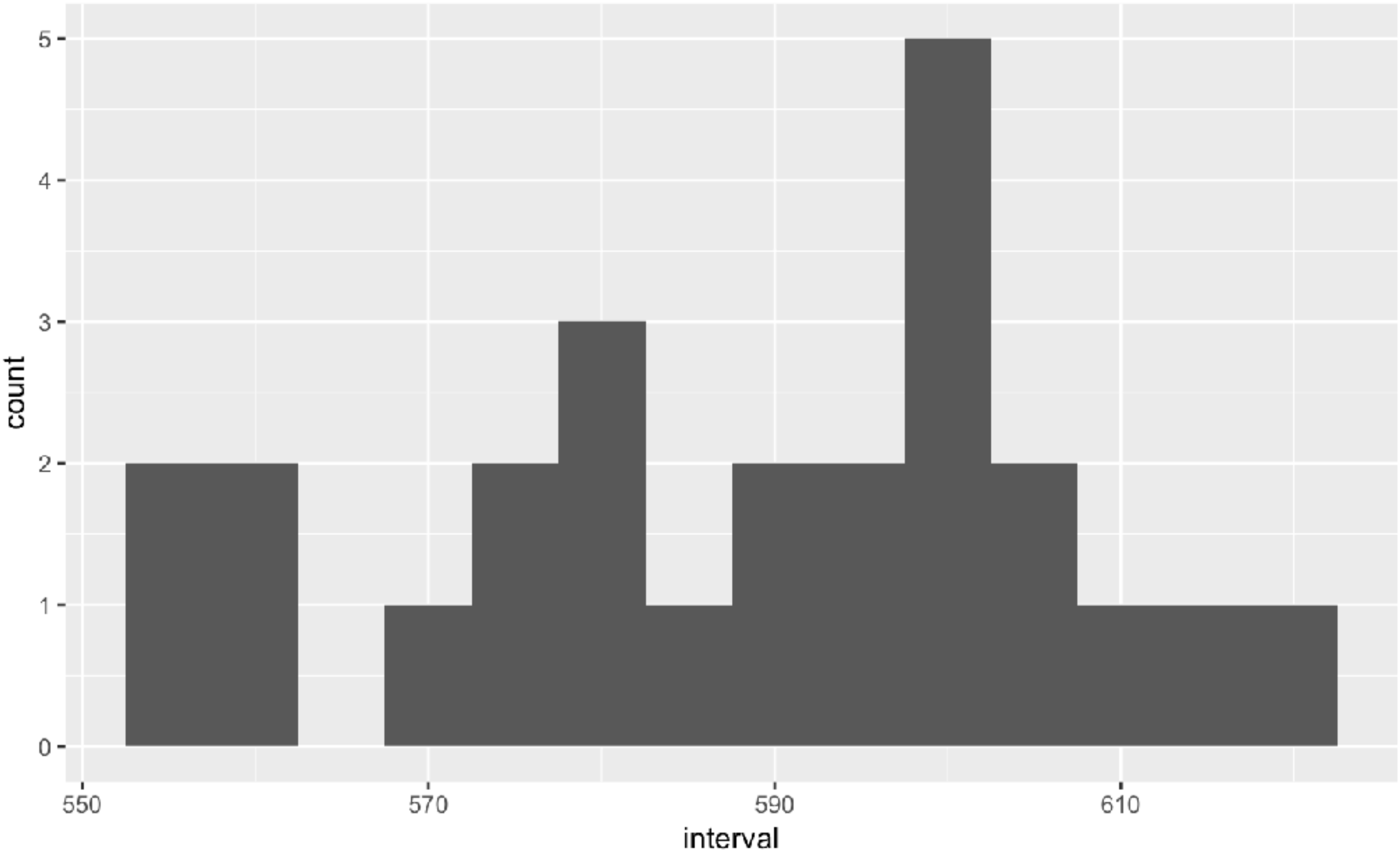


Image Analysis for Pulmonary Emboli

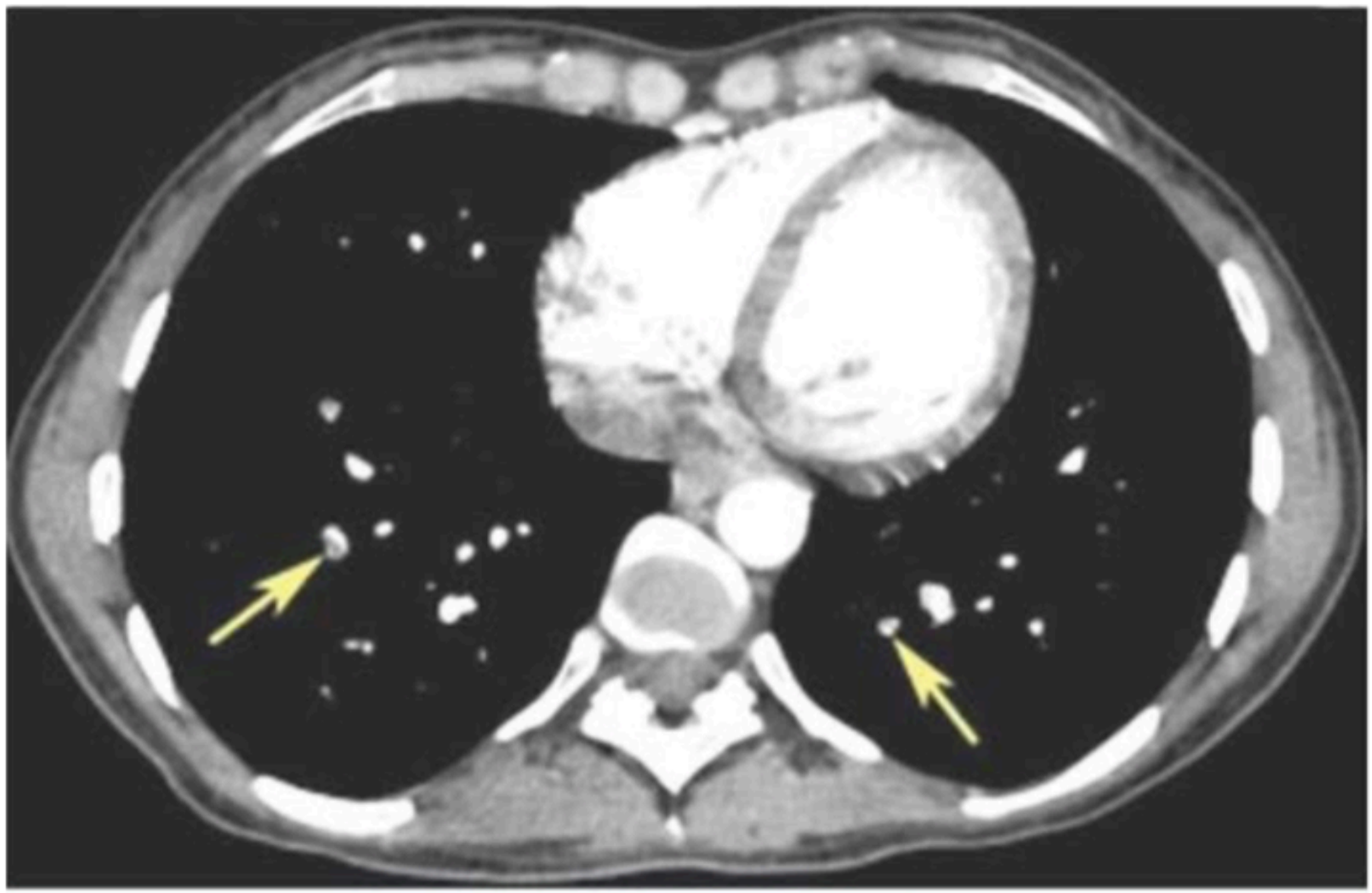
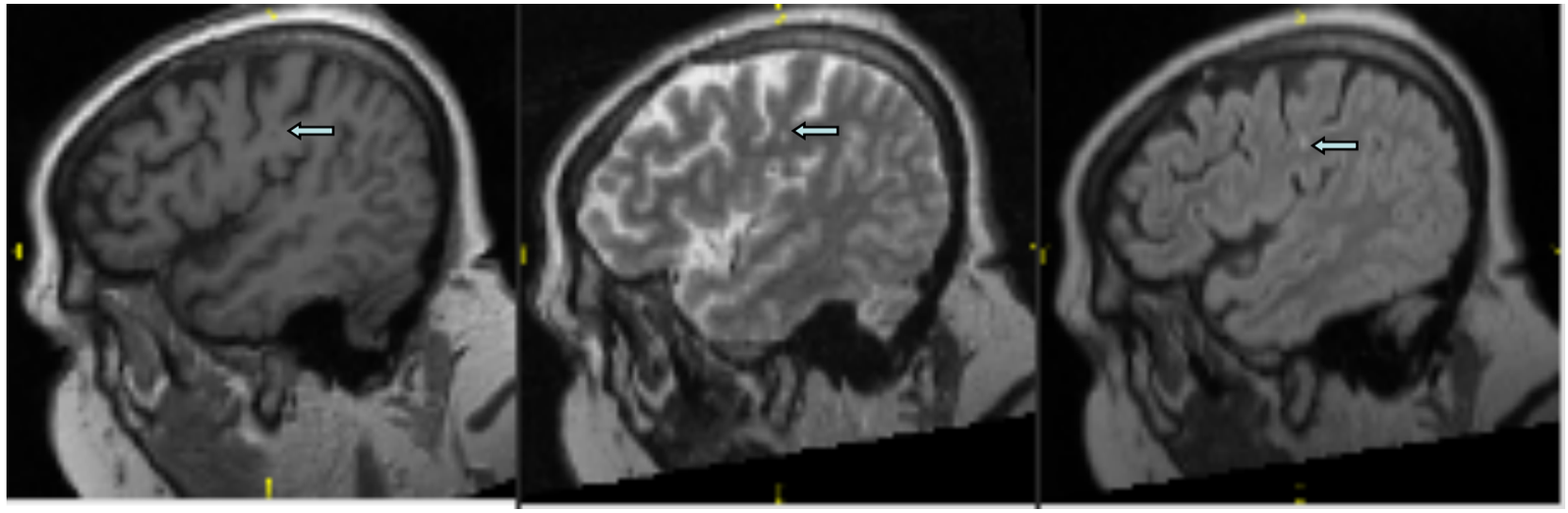


Figure 2: Bilateral pulmonary emboli in third and fourth order branch points of the pulmonary arteries. Small emboli distally located can be a diagnostic challenge.

Lupus Lesions

- Automatic Analysis of White Matter Abnormalities in Neuropsychiatric SLE (Lupus)
- About 1.5 Million Americans with Lupus, Underlying Pathologic Processes Unknown – Possibly Vascular



Hypointense on T1 | Hyperintense T2 | Hyperintense on FLAIR

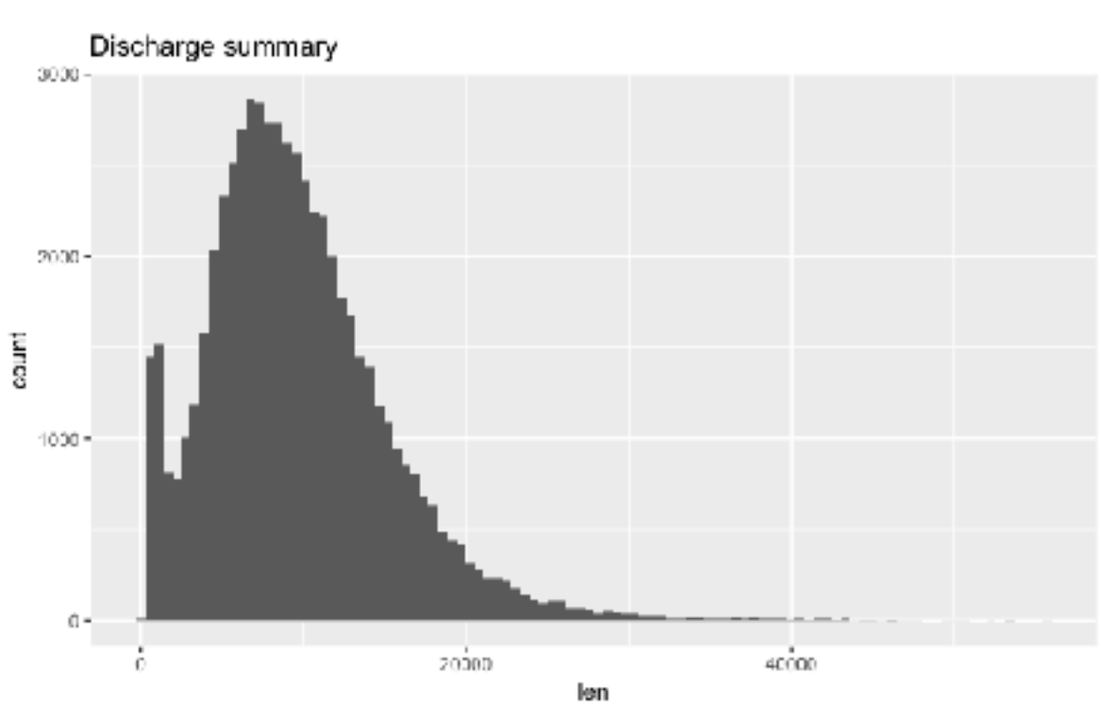
Images: Bockholt et al

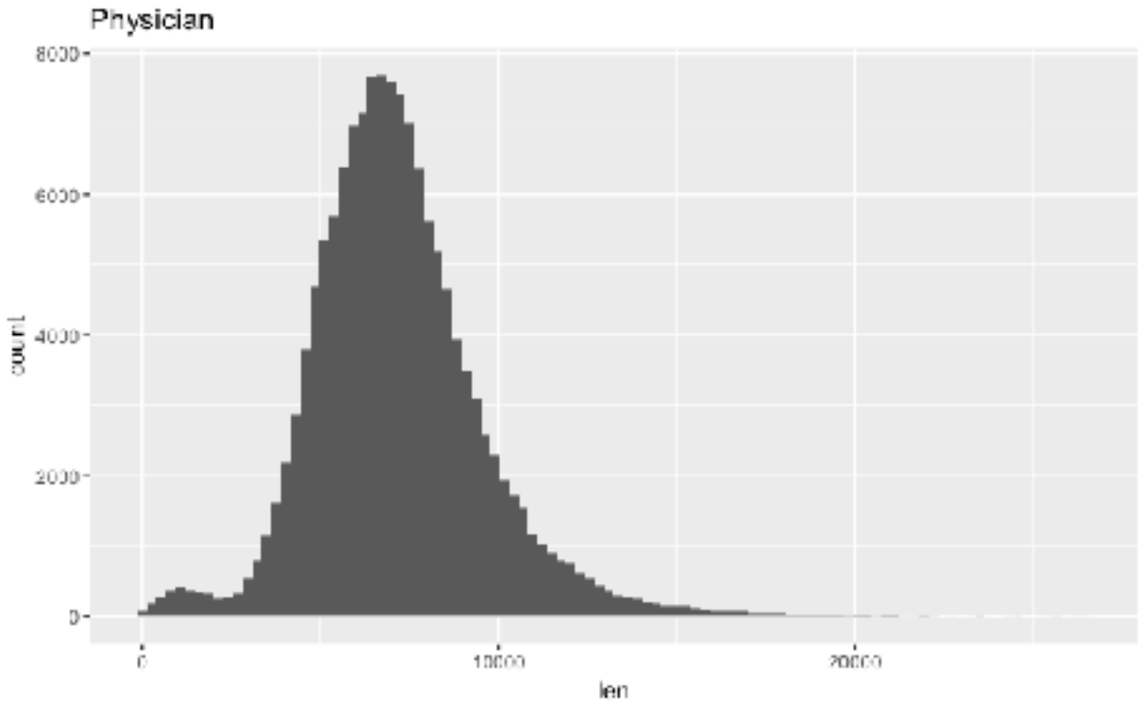
Slide from Ron Kikinis

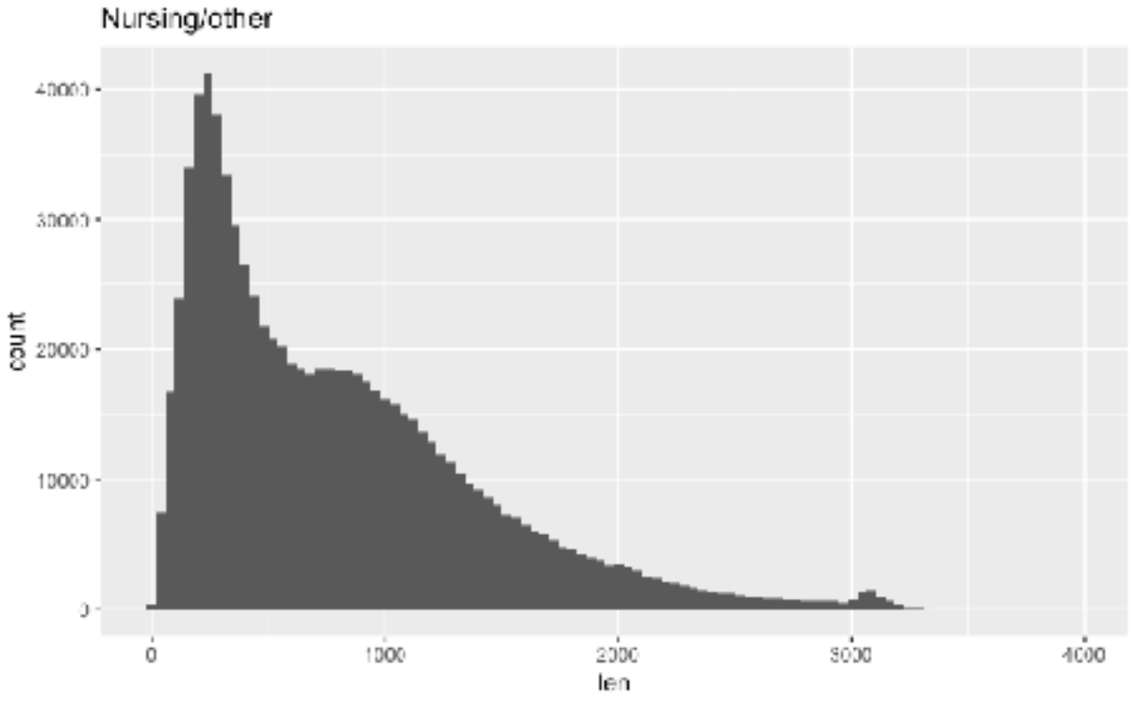
Clinical Notes in MIMIC

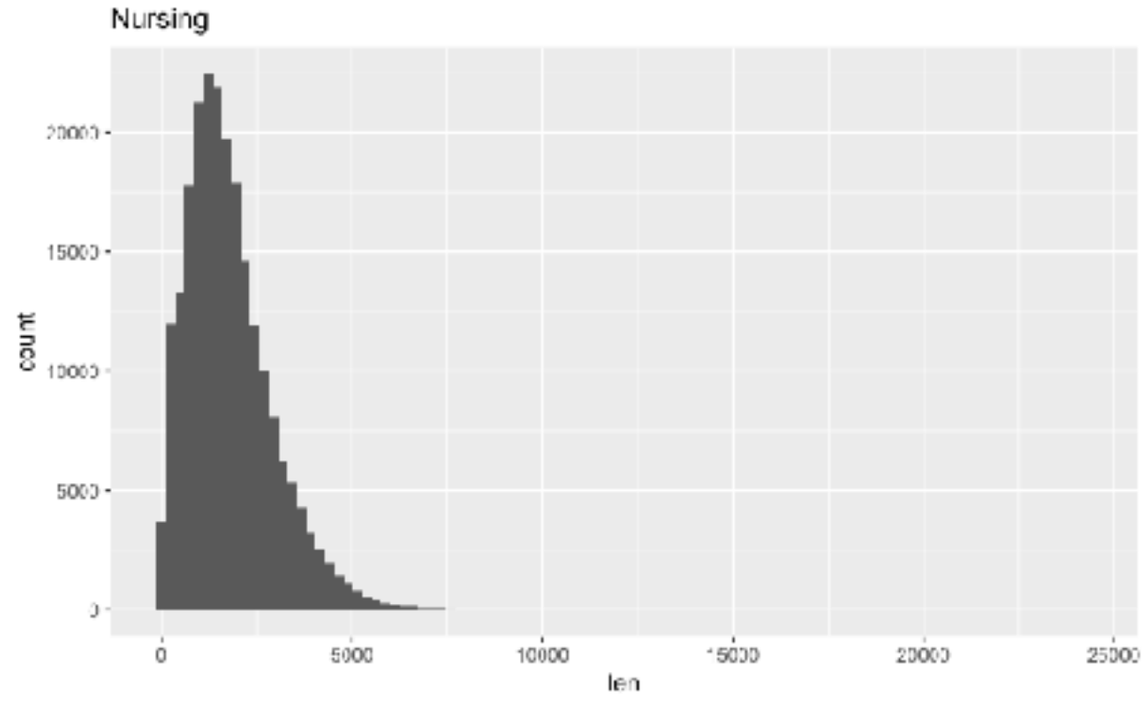
Nursing/other	822497
Radiology	522279
Nursing	223556
ECG	209051
Physician	141624
Discharge summary	59652
Echo	45794
Respiratory	31739
Nutrition	9418
General	8301
Rehab Services	5431
Social Work	2670
Case Management	967
Pharmacy	103
Consult	98

Lengths of different kinds of notes









A brief nursing note

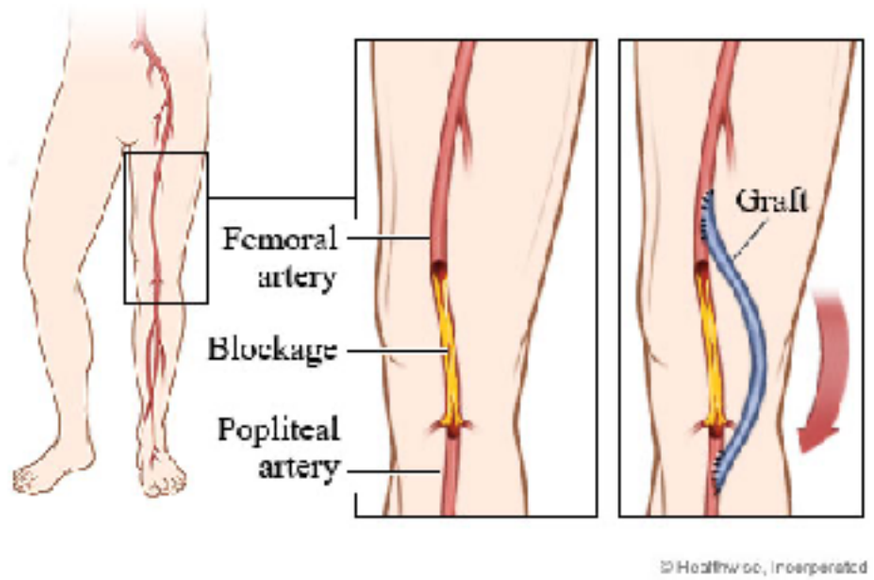
Hypotension (not Shock)
Assessment:
Pt remains on phenylephrine drip at 0.75 mcg/kg/min
Action:
No titration needed at this time
Response:
BP stable at > 100, MAP >65
Plan:
Wean Neo if tolerated
Wound infection
Assessment:
Anterior groin area open and oozing mod amts thin pink tinged serous fluid
Pt stooling, with small amts stool on dsg and dangerously close to open wound
Action:
Urology resident in to change dressing
Propofol increased to 100 mcg nad fentanyl 100 mcg given for comfort during dsg change
Flexiseal inserted to help contain bowel movements
Stool sent for c diff.
Response:
Pt comfortable during proceedure
Plan:
Continue sedation as needed, increasing Propofol to 100 mcg for sedation during dsg changes.
Keep wound area as clean as possible, check for incontinence of stool as needed

Admission Date: [**2198-7-16**] Discharge Date: [**2198-7-28**]

Date of Birth: [**2153-5-26**] Sex: F

Service: SURGERY

Discharge Summary



Allergies:
No Known Allergies / Adverse Drug Reactions

Attending:[**First Name3 (LF) 1234**]
Chief Complaint:
Leg pain, erythema and swelling secondary to infection of left femoral-poplital bypass

Major Surgical or Invasive Procedure:
1. Incision and drainage and pulse irrigation of left groin and left above-knee popliteal site incisions with xxploration of bypass graft ([**2198-7-16**])
2. Excision of entire left common femoral artery-to-above-knee popliteal artery bypass graft; Repair of common femoral artery and above-knee popliteal artery with harvested left arm cephalic vein ([**2198-7-18**])
3. I and D/washout of left groin with complex wound closure over 2 drains

History of Present Illness:
Ms. [**Known lastname **] is a 45 y/o F who underwent a left fem-AK [**Doctor Last Name **] BPG with PTFE over one month ago on [**2198-6-11**]. She had been doing well postoperatively, and was seen in the clinic 6 days prior to presentation. At this time, she acutely developed nausea/vomiting, fevers, and progressive redness/swelling/pain of her left thigh
directly at the surgical incision. She has been unable to keep down food or liquids. At the time, she denied any ischemic-type pain in her lower leg, and denied any chest pain or shortness of breath.

Past Medical History:
PMH: current smoker (1-PPD), cocaine abuse (ceased 6-months prior), asthma, diabetes type 2

PSH: bilateral lower extremity angiograms ([**2198-5-10**]), L knee surgery x2, appendectomy, tonsillectomy, L fem-AK [**Doctor Last Name **] [**2198-6-11**]

Social History:

Family History:
Mother had an abdominal aortic aneurysm status post repair, MI in her mid 50s, carotid stenosis, cervical cancer, coronary artery disease, other vascular lesions which were stented. She died due to complications of a procedure. The patient's father died young. The patient has one cousin with cervical cancer. Her maternal grandmother had an MI in her 60s. Maternal grandfather with MI, hypertension, and hypercholesteremia.

Physical Exam:
Upon presentation,

Vital Signs: Temp: 101.9 RR: 16 Pulse: 98 BP: 114/62
Neuro/Psych: Oriented x3, Affect Normal.
Neck: No masses, Trachea midline, No right carotid bruit, No left
carotid bruit.
Nodes: No clavicular/cervical adenopathy, No inguinal adenopathy.

Skin: Abnormal: Cellulitis L thigh.
Heart: Regular rate and rhythm.
Lungs: Clear, Normal respiratory effort.
Gastrointestinal: Non distended, No hepatosplenomegally, No hernia, No AAA.
Rectal: Not Examined.
Extremities: No popiteal aneurysm, No femoral bruit/thrill, No RLE edema, No LLE Edema, No varicosities, abnormal: Tenderness, erythema of L thigh.

Pulse Exam (P=Palpation, D=Dopplerable, N=None)
RUE Radial: P.
LUE Radial: P.
RLE Femoral: P. Popiteal: P. DP: P. PT: P.
LLE Femoral: P. Popiteal: P. DP: P. PT: P. Other: Graft: palp.

DESCRIPTION OF WOUND: R thigh incision without breakdown, but tender, erythematous, and swollen especially superiorly. No evidence of drainage or underlying fluctuance. pulses all palpable

Pertinent Results:
[**2198-7-16**] 02:30AM BLOOD WBC-7.4 RBC-3.21*# Hgb-9.6* Hct-27.6*
MCV-86 MCH-29.9 MCHC-34.7 RDW-13.6 Plt Ct-161
[**2198-7-19**] 05:54AM BLOOD WBC-5.6 RBC-3.32* Hgb-10.2* Hct-28.7*
MCV-87 MCH-30.7 MCHC-35.5* RDW-14.1 Plt Ct-184
[**2198-7-27**] 05:06AM BLOOD WBC-7.8 RBC-2.98* Hgb-9.0* Hct-26.9*
MCV-90 MCH-30.4 MCHC-33.6 RDW-15.8* Plt Ct-398

[**2198-7-16**] 09:05AM BLOOD PT-13.9* PTT-36.4* INR(PT)-1.2*
[**2198-7-19**] 05:54AM BLOOD PT-13.1 PTT-29.8 INR(PT)-1.1

[**2198-7-16**] 02:30AM BLOOD Glucose-177* UreaN-20 Creat-1.0 Na-135
K-3.7 Cl-99 HCO3-23 AnGap-17
[**2198-7-27**] 05:06AM BLOOD Glucose-72 UreaN-10 Creat-0.6 Na-142
K-3.5 Cl-108 HCO3-27 AnGap-11
[**2198-7-17**] 10:15PM BLOOD CK(CPK)-99
[**2198-7-27**] 05:06AM BLOOD Calcium-8.3* Phos-4.6* Mg-1.9

[**2198-7-16**] 02:30AM URINE Blood-TR Nitrite-NEG Protein-100
Glucose-NEG Ketone-NEG Bilirub-NEG Urobiln-NEG pH-5.5 Leuks-SM

Blood Culture, Routine (Final [**2198-7-22**]):   NO GROWTH.

Brief Hospital Course:
The patient was admitted to the surgery service for evaluation and treatment of her lower extremity bypass graft infection.

Neuro: The patient received IV pain medications with good effect and adequate pain control. When tolerating oral intake, the patient was transitioned to oral pain medications.

CV: The patient was stable from a cardiovascular standpoint; vital signs were routinely monitored.

Pulmonary: The patient was stable from a pulmonary standpoint; vital signs were routinely monitored. Good pulmonary toilet, early ambulation and incentive spirometry were encouraged throughout this hospitalization.

GI/GU/FEN:
Post operatively, the patient was made NPO with IVF.
The patient's diet was advanced when appropriate, which was tolerated well.
The patient's intake and output were closely monitored, and IVF were adjusted when necessary. The patient's electrolytes were routinely followed during this hospitalization, and repleted when necessary.

ID: The patient's white blood count and fever curves were closely watched for signs of infection. Final blood cultures were negative.

Endocrine: The patient's blood sugar was monitored throughout this stay; insulin dosing was adjusted accordingly and kept within normal range.

Hematology: The patient's complete blood count was examined routinely; no transfusions were required during this stay.

Prophylaxis: The patient received subcutaneous heparin during this stay, and was encouraged to get up and ambulate as early as possible.

At the time of discharge, the patient was doing well, afebrile with stable vital signs. The patient was tolerating a regular diet, ambulating, voiding without assistance, and pain was well controlled.

Medications on Admission:
Albuterol INH PRN
Fexofenadine 60mg 1 tablet [**Hospital1 **]
Fluticasone 50mcg two puffs daily
Percocet PRN
Glargine 35 units
Humalog SS
Lisinopril 40mg qd
Crestor 40mg qd
Metformin 1000mg [**Hospital1 **]
Reglan 5QACHS
Protonix 40mg qd
Tizanidine 4PRN
ASA 81mg qd

Discharge Medications:
1. albuterol sulfate 90 mcg/Actuation HFA Aerosol Inhaler Sig: Two (2) Puff Inhalation Q4H (every 4 hours) as needed for wheeze.
2. fexofenadine 60 mg Tablet Sig: One (1) Tablet PO BID (2 times a day).
3. fluticasone-salmeterol 100-50 mcg/dose Disk with Device Sig: One (1) Disk with Device Inhalation [**Hospital1 **] (2 times a day).
4. rosuvastatin 20 mg Tablet Sig: Two (2) Tablet PO DAILY (Daily).
5. aspirin 81 mg Tablet, Chewable Sig: One (1) Tablet, Chewable PO DAILY (Daily).
6. docusate sodium 100 mg Capsule Sig: One (1) Capsule PO BID (2 times a day).
7. hydromorphone 2 mg Tablet Sig: One (1) Tablet PO Q4H (every 4 hours) as needed for pain.
Disp:*30 Tablet(s)* Refills:*0*
8. pantoprazole 40 mg Tablet, Delayed Release (E.C.) Sig: One (1) Tablet, Delayed Release (E.C.) PO Q24H (every 24 hours).
9. dicloxacillin 500 mg Capsule Sig: One (1) Capsule PO Q6H (every 6 hours) for 4 weeks: Take 1 tablet every 4 hours for a total 4 week course. First day was [**7-27**].
10. lisinopril 40 mg Tablet Sig: One (1) Tablet PO DAILY (Daily).
11. cilostazol 100 mg Tablet Sig: One (1) Tablet PO BID (2 times a day)
12. Metformin 1000mg [**Hospital1 **]
13. Humalog SS
14. Glargin 35 units (at discretion of patient while monitoring blood sugars, to be followed-up by PCP)

Discharge Disposition:
Home With Service

Facility:
[**Hospital 119**] Homecare

Discharge Diagnosis:
Infected left femoral-popliteal bypass graft

Discharge Condition:
Mental Status: Clear and coherent.
Level of Consciousness: Alert and interactive.
Activity Status: Ambulatory - Independent

Discharge Instructions:
What activities you can and cannot do:
?????? When you go home, you may walk and go up and down stairs with an ace-wrap or compression stocking on your left leg.
?????? You may shower (let the soapy water run over groin incision, rinse and pat dry)
?????? Your incision may be left uncovered, unless you have small amounts of drainage from the wound, then place a dry dressing or band aid over the area that is draining, as needed
-Monitor drainage from both JP drains. If either drains less than 20cc in one day, please call Dr. [**Last Name (STitle) 2866**] at his clinic (see number below). Your visiting nurse will teach you how to monitor and care for your drains.
?????? No heavy lifting, pushing or pulling (greater than 5 lbs) for 1 week (to allow groin puncture to heal)
?????? After 1 week, you may resume sexual activity
?????? After 1 week, gradually increase your activities and distance walked as you can tolerate
?????? No driving until you are no longer taking pain medications
?????? Call and schedule an appointment to be seen in [**4-6**] weeks for post procedure check and ultrasound

What to report to office:
?????? Numbness, coldness or pain in lower extremities
?????? Temperature greater than 101.5F for 24 hours
?????? New or increased drainage from incision or white, yellow or
green drainage from incisions
?????? Bleeding from groin puncture site

SUDDEN, SEVERE BLEEDING OR SWELLING (Groin puncture site)
?????? Lie down, keep leg straight and have someone apply firm
pressure to area for 10 minutes. If bleeding stops, call
vascular office [**Telephone/Fax (1) 1237**]. If bleeding does not stop, call
911 for transfer to closest Emergency Room.

Followup Instructions:
Please follow-up with Dr. [**Last Name (STitle) **] on [**8-7**]: call his clinic at
([**Telephone/Fax (1) 2867**] to schedule an appointment.

Please follow-up with Dr. [**Last Name (STitle) 2866**] in two weeks; call his clinic
at ([**Telephone/Fax (1) 2868**] to schedule an appointment.

Completed by:[**2198-7-31**]

Data Standards

- OHDSI
- FHIR (Fast Healthcare Interoperability Resources) — pronounced “fire”
- HL7

```
<Patient xmlns="http://hl7.org/fhir">
  <id value="glossy"/>
  <meta>
    <lastUpdated value="2014-11-13T11:41:00+11:00"/>
  </meta>
  <text>
    <status value="generated"/>
    <div xmlns="http://www.w3.org/1999/xhtml">
      <p>Henry Levin the 7th</p>
      <p>MRN: 123456. Male, 24-Sept 1932</p>
    </div>
  </text>
  <extension url="http://example.org/StructureDefinition/trials">
    <valueCode value="renal"/>
  </extension>
  <identifier>
    <use value="usual"/>
    <type>
      <coding>
        <system value="http://hl7.org/fhir/v2/0203"/>
        <code value="MR"/>
      </coding>
    </type>
    <system value="http://www.goodhealth.org/identifiers/mrn"/>
    <value value="123456"/>
  </identifier>
  <active value="true"/>
  <name>
    <family value="Levin"/>
    <given value="Henry"/>
    <suffix value="The 7th"/>
  </name>
  <gender value="male"/>
  <birthDate value="1932-09-24"/>
  <careProvider>
    <reference value="Organization/2"/>
    <display value="Good Health Clinic"/>
  </careProvider>
</Patient>
```

Resource Identity & Metadata

Human Readable Summary

Extension with URL to definition

Standard Data:

- MRN
- Name
- Gender
- Birth Date
- Provider

Terminology Standards

- LOINC
- NDC
- ICD-9, ICD-10
- SNOMED
- DSM-5
- ...

- all gathered in the UMLS Metathesaurus
  - https://uts.nlm.nih.gov/home.html

Take-away lessons

- Know your data!
- “Harmonization” is difficult and time-consuming
- Standards are often lacking

- Can tons of observational data approximate the results of RCTs?