What is Intracerebral Hemorrhage?

- Subtype of stroke
- Occurs when a diseased blood vessel within the brain bursts, allowing blood to leak inside the brain.
- The sudden increase in pressure within the brain can cause damage to the brain cells surrounding the blood.

Why is it important to study ICH?

- ICH is a very serious condition and the most devastating type of stroke with a high rate of mortality

- Studies show that the overall 30-day mortality is 34.6%. After six months the mortality rate increases to 38.6% [Godoy, 2006]

Why is it important to study ICH?

- Yet, ICH is not treated as a main stream condition which leads to a lack of full understanding of its mechanisms and absence of commonly used guidelines [ICHNewInsights, 2016]

- In particular, very few studies focus on identifying factors that affect the hospital length of stay of a patient

ICH Length of Stay (LoS)

- A study by Chan et al. in 2014 defined prolonged LoS for ICH patients to be equal or greater than 10 days [Biomed Res Int, 2014]
  - Pop. 1599 patients, 436 prolonged, 1163 not prolonged (Taiwan National Health Insurance Research Database, NHIRD)

- But it did not explain patterns of patients that stayed longer at the hospital

ICH Length of Stay (LoS)

- Kim et al. in 2013 published a general study of patients from the Discharge Injury Survey of the Korea Centers for Disease Control and Prevention
  - 17,364 cases of stroke from 2005 to 2008
- The study shows that ICH patients (n=3250, 18.7%) have an average length of stay of 28.9 days
- Patients who underwent surgery had longer LoS
- “The emergency route of admission and other diagnosis increased LoS, whereas hypertension and diabetic mellitus reduced LoS.”

This study

What are the patterns of patients which stay longer at the ICU?

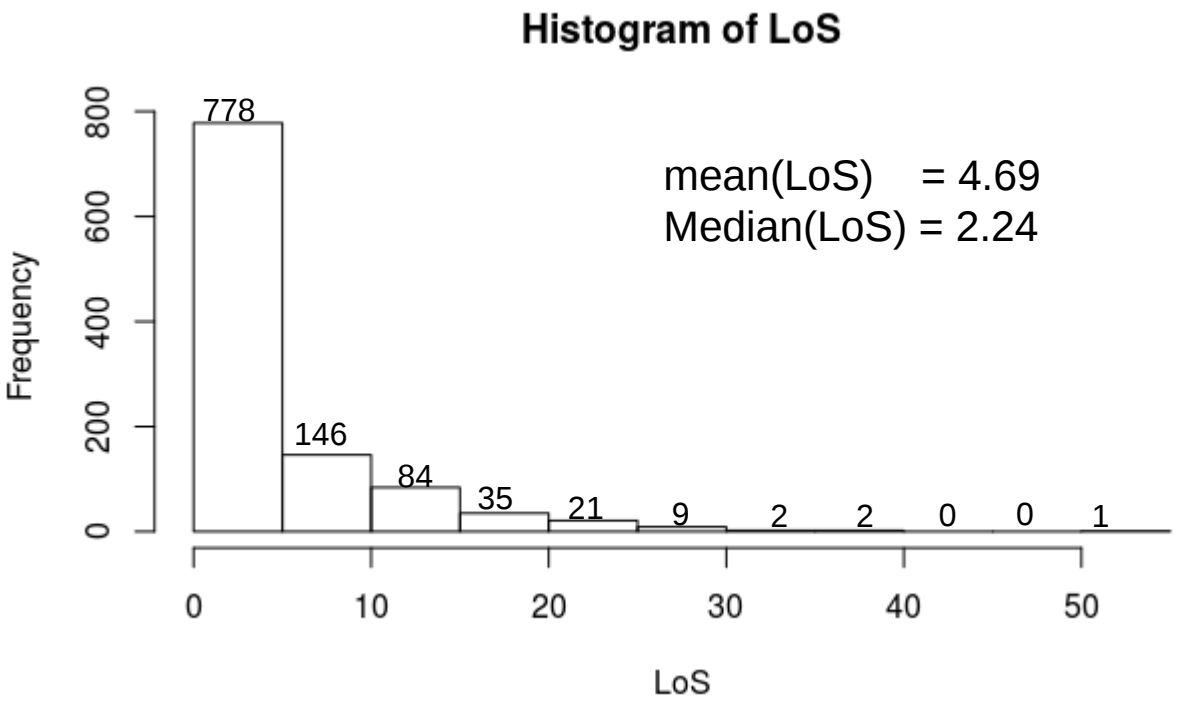
This study

- An answer to this question may:
  - Help providing better and timely health care to patients
  - Help improving resources allocation (people, medications, beds, instruments etc)
  - Help finding inconsistencies in procedures or treatments

MIMIC-III ICH ICU Length of Stay

Case study:
From MIMIC-III database
Intracerebral Hemorrhage
(ICD9: 431 or 432.9)

n=1078



Gender	M	53.6	57.5	56.0	63.0	76.0	67.0	0	1	0	593
	F	46.4	42.5	44.0	37.0	24.0	33.0	2	1	1	485
Age	>50	90%	88%	89%	80%	90%	78%	1	2	1	964
	<=50	10%	12%	11%	20%	10%	22%	1	0	0	114

Deaths

Up to 1 day of admission	In hospital
86 (8%) (known DoD)	356 (33%)

Challenge

- How to study LoS using 26 tables and 500+ heterogeneous variables?

- Each ICU is represented by:
  - Static data (demographics, insurance etc)
  - Series of temporal events
    - Non periodic
    - Non fixed-length

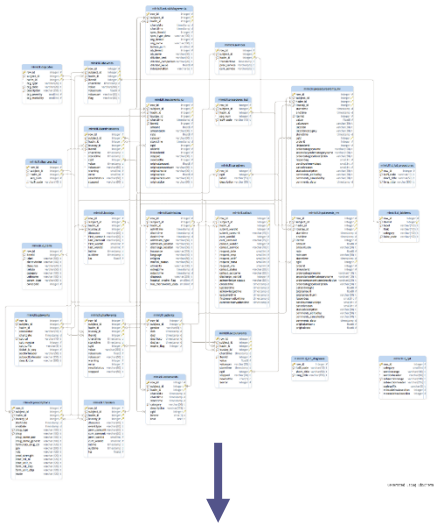
Challenge

- Represent data and study patterns for LoS using different methods, compare them

- Two tasks:
  - Exploratory analysis and Prediction (prediction will be approached later, not in these slides)

Data and Knowledge Representation

Propositional representation



G	A	L	It1	It2	Itn
M	51	2.1	...	...	...
F	76	1.2	...	...	...
M	68	10.3	...	...	...

G: Gender, A: Age, L: LoS

Propositional representation

- Static variables: easy
  - Select the ones already available in the dataset
- Temporal events?
  - Various options
  - Simpler:
    - Select best evidence from the literature and represent items as aggregates

Choice of variables

- From the literature
  - Age, Gender, Ethnicity, Alcoholism, hypertension
  - Other findings:
    - Anticoagulants (warfarin, heparin)
    - Constipation
    - Urine amount
  - Also related, but not yet used:
    - Smoking habits, hypocholesterolemia, use of drugs, amyloid angiopathy and cerebral microbleeds, volume and growth of the hematoma, depression, infection, malnutrition, dehydration, glucose levels, seizure control

Propositional Representation

- Temporal events
  - Alcoholism
  - Blood pressure
  - Hypertension
  - Platelets count
  - Platelets intake
  - Prothrombin time
  - Stool
  - Urine
  - Warfarin, heparin
- Static variables
  - Age
  - Gender
  - Ethnicity
  - Admission Type
  - Admission Location
  - First care unit
  - First ward ID
  - Insurance
  - Language
  - Religion
  - Marital Status

Outcome: LoS

Propositional Representation

- Flat table
  - Columns for demographics
  - One column for each item
    - Values: rates along the time
      - total occurrences (frequency) or
      - amounts (dosage or collection)
- Dimension: 1078 ICU cases (rows)
  26 variables (columns)
- Outcome: LoS

Propositional Representation

Populating Table

- Temporal constraint:
  - Events to study are collected up to one day after hospital admission

Propositional table populated

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max	NA's or zeros
Age	24.240	59.450	71.900	80.090	81.210	300.210	0
Hypert	1.131	6.115	10.423	17.545	17.950	288.000	782
Platelet Intake	1149	12714	22478	52836	55807	902352	1024
ProThrombine Time	10.600	13.100	15.400	48.300	96.650	150.000	693
Urine	35.300	1681.300	3172.200	8924.400	7162.800	766560.000	520
Stool	225.500	287.600	639.700	1247.6	1521.300	3995.2	1072
Syst BP	86.210	127.110	137.270	135.300	143.690	175.000	906
Dias BP	26.000	57.410	64.540	64.360	70.480	88.750	906
BP Alarm Low	70.000	90.000	90.000	94.650	95.000	160.000	917
BP Alarm High	90.000	150.000	160.000	156.000	162.500	215.000	917
Platelet Count	1.000	155.200	409.500	412.800	659.800	912.000	0
LoS	0.112	1.185	2.240	4.686	5.760	51.795	0

Numerical variables

Propositional table populated

Categorical variables

Insurance	'Government'	29
	'Medicaid	68
	'Medicare'	674
	'Private'	286
	'Self Pay'	21
Language	?	401
	'ASIAN'	28
	'ENGL'	578
	'EURO'	26
	'SPAN'	31
	'OTHER'	14
Religion	'CATHOLIC'	314
	'NOT SPECIFIED'	231
	'UNOBTAINABLE'	228
	'PROTESTANT QUAKER'	117
	'JEWISH'	73
	'OTHER'	47
	(Other)	68
Marital Status	?	94
	'DIVORCED'	58
	'MARRIED'	551
	'SEPARATED'	10
	'SINGLE'	188
	'UNKNOWN (DEFAULT)'	7
	'WIDOWED'	170

Gender	F	485
	M	593
Ethnicity	'WHITE'	783
	'UNKNOWN/NOT SPECIFIED'	81
	'BLACK'	78
	'ASIAN'	43
	'HISPANIC/LATINO'	40
	'OTHER'	35
	(Other)	18
Admission Type	EMERGENCY	1066
	URGENT	12
Admission Loc	'CLINIC REFERRAL/PREMATURE'	261
	'EMERGENCY ROOM ADMIT'	751
	'PHYS REFERRAL/NORMAL DELI'	4
	'TRANSFER FROM HOSP/EXTRAM'	57
	'TRANSFER FROM OTHER HEALT'	3
	'TRANSFER FROM SKILLED NUR'	2
1st Care Unit	'CCU'	45
	'CSRU'	14
	'MICU'	165
	'SICU'	654
	'TSICU'	200
1st Ward ID	7	21
	12	53
	14	207
	15	49
	22	1
	23	96
	33	317
	38	2
	50	17
	52	10
	57	305

Experiments

Experiments

- Separated ICU stays in two groups

  >= 10 ICU days (pos=154)

  < 10 ICU days (neg=924)

  (chosen using change point analysis)

Experiments

- Exploratory analysis (full data)
  - Stepwise Logistic Regression (propositional)
    - Step and glm R functions

Results
Exploratory Analysis

Propositional

- Logistic regression (reference: long stay)

Variable	Coefficient	Stderr	p-value	Sign.	CI (2.5-97.5%)
(Intercept)	0.78036	0.36250	0.03134	*	(0.03519 , 1.42506)
Age	0.01666	0.00521	0.00138	**	(0.00775 , 0.02759)
Stool	-0.00178	0.00059	0.00241	**	(-0.00341 ,-0.00073)
Syst. BP	-0.00555	0.00152	0.00026	***	(-0.00848 ,-0.00252)

- Only using significant variables given by the last step of glm
- Age is positively correlated with LoS
  - An increase of 1 year increases the probability of staying in the ICU ~0.02 day
- Stool and Syst. BP are negatively correlated with LoS
  - An increase of 1 ml of stool may reduce the probability of staying in the ICU in ~0.002 days
  - An increase of 1 mmHg may reduce the probability of staying in the ICU in ~0.005 days

Main findings

- Age, BP and Stool may affect LoS
  - Old age was already mentioned in the literature as well as Syst. BP