DAT-ICU
Team #8 - Data Quality

January 20-21, 2018
#DATICU2018

The survivors !

Patricia SERRE, Data quality manager, AP-HP (leader)

Audrey MOUNGNY, data scientist, AP-HP (presenter)

Dr Marc WYSOCKI, MD, Product Manager, GE Healthcare Digital

Dr Christel DANIEL, MD, Medical informatics, AP-HP

Thanks to:

Dr Jonas POCHARD, intensivist, biostatistician, AP-HP

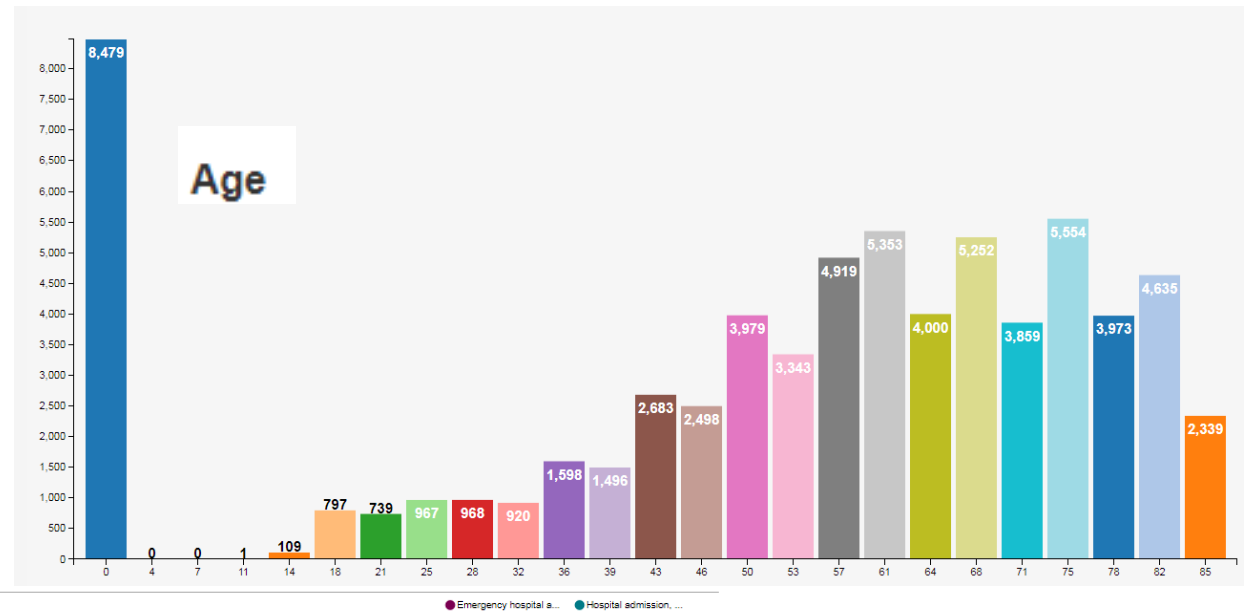**Kankoe SALLAH**, data scientist, AP-HP

Thanks to their efforts yesterday!

Massachusetts Institute of Technology · srlf SOCIÉTÉ DE RÉANIMATION DE LANGUE FRANÇAISE · UNIVERSITÉ PARIS DESCARTES · ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS · www.aphp.fr

- **Characterizing the MIMIC data and its quality**
  - ▸ What does the MIMIC data look like ? (characterization)

  - ▸ What is the quality of the data? (data quality assessment)
    - *Can we define quality indicators for strategic data*
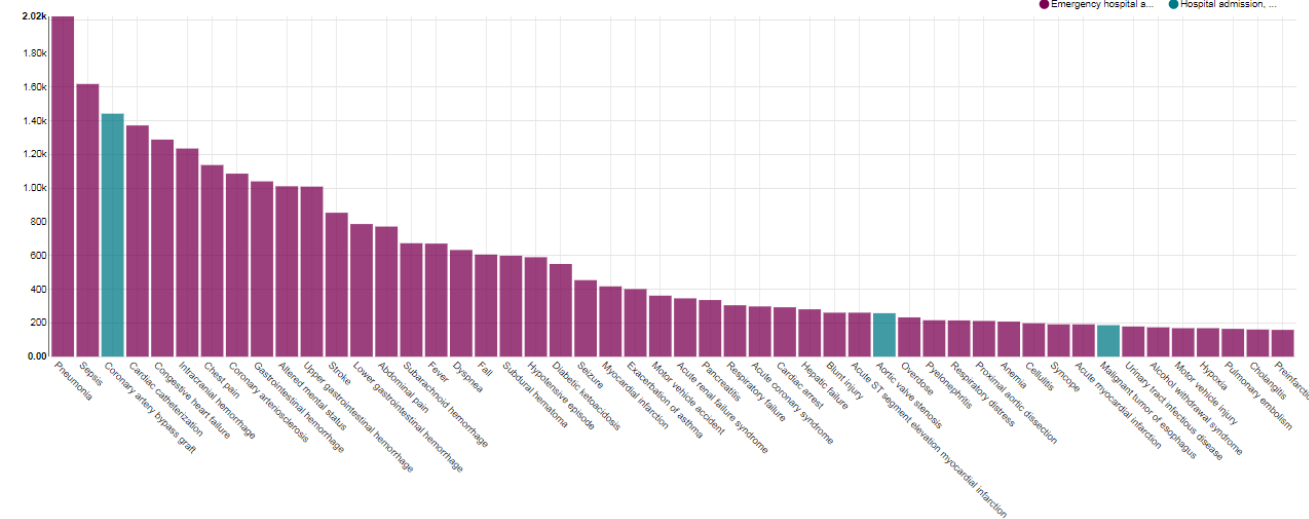 - ▸ How to report on data quality(visualization)
    - *User friendly diagrams*

46480 patients

57790 admissions

Top 20 procedures

Concept	id	nb
Subsequent hospital visit by physician	4027135	147067
Patient evaluation and management	4033552	106490
Evaluation and management of inpatient	4313303	105559
Inpatient care	4214577	102331
Peripheral venous cannula insertion	4049832	60347
Plain chest X-ray	4163872	31953
Central venous cannula insertion	4052413	15357
Venous catheterization, not elsewhere classified	2002174	14731
Insertion of catheter into artery	4213288	14314
Electrocardiographic procedure	4163951	13664

Massachusetts Institute of Technology

srlf SOCIÉTÉ DE RÉANIMATION DE LANGUE FRANÇAISE

UNIVERSITÉ PARIS DESCARTES

ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS

www.aphp.fr

Compliance

▸ Impossible negative values – e.g intracranial pressure

No matching concept	497k	60.9%
Base excess of blood	198k	24.2%
Level of arousal	107k	13.1%
Urine output	6.22k	0.762%
Central venous pressure (CVP)	4.15k	0.509%
Intracranial pressure	2.61k	0.320%
Mean blood pressure	1.24k	0.152%
Pulmonary artery Diastolic blood pressure	45.0	0.00552%
Pulmonary artery Mean blood pressure	37.0	0.00454%

400 mesures < -10

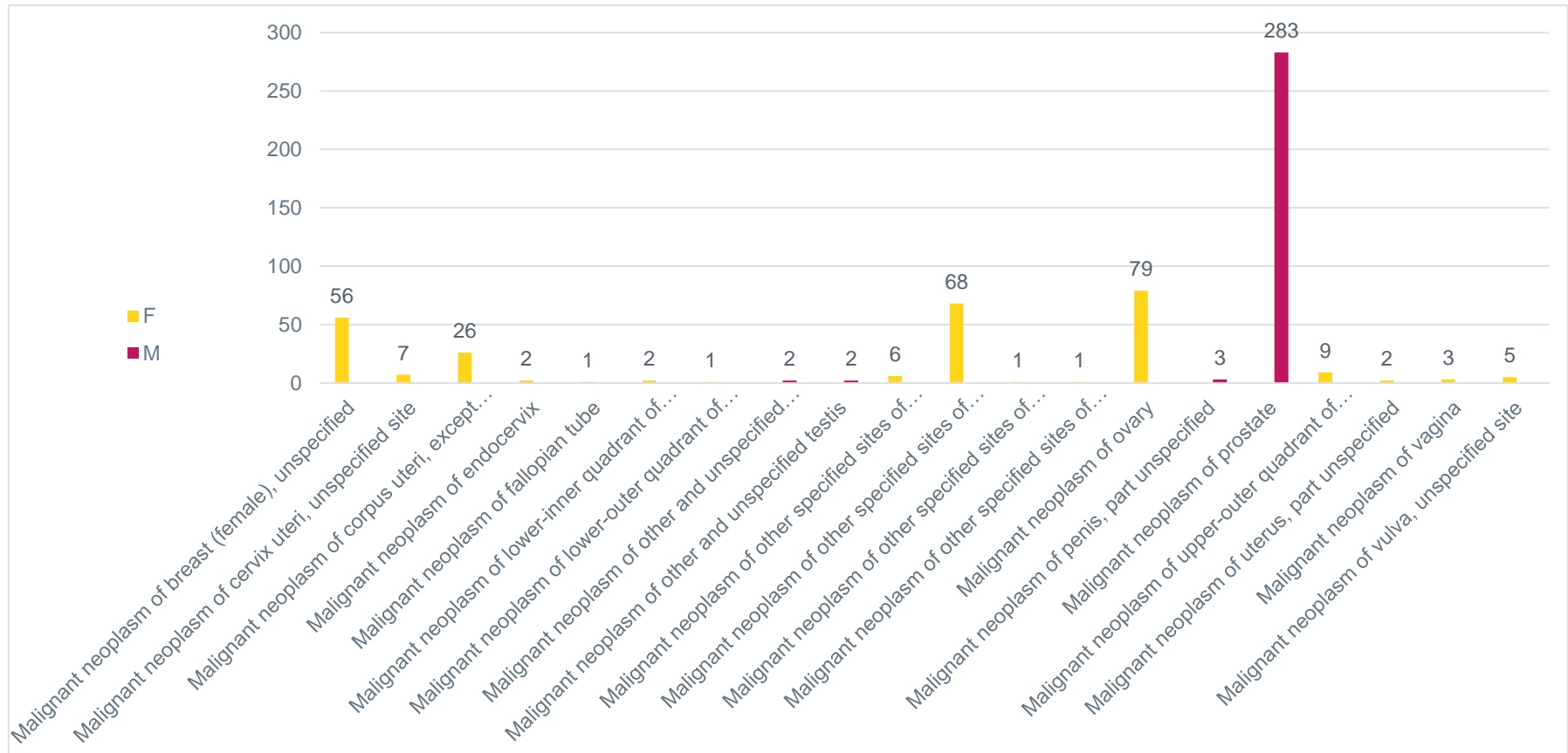▸ **44 374** date of birth > 2018

Completeness

- **178694 visit detail occurence without related visit occurrence ID (required field)**

Plausibility
How many females with prostate cancer ?/males with uterus cancer?

Bar chart showing counts of cancer diagnoses by sex (F in yellow, M in magenta):

- Malignant neoplasm of breast (female), unspecified: 56
- Malignant neoplasm of cervix uteri, unspecified site: 7
- Malignant neoplasm of corpus uteri, except...: 26
- Malignant neoplasm of endocervix: 2
- Malignant neoplasm of fallopian tube: 1
- Malignant neoplasm of lower-inner quadrant of...: 2
- Malignant neoplasm of lower-outer quadrant of...: 1
- Malignant neoplasm of other and unspecified...: 2
- Malignant neoplasm of other and unspecified testis: 2
- Malignant neoplasm of other specified sites of...: 6
- Malignant neoplasm of other specified sites of...: 68
- Malignant neoplasm of other specified sites of...: 1
- Malignant neoplasm of other specified sites of...: 1
- Malignant neoplasm of ovary: 79
- Malignant neoplasm of penis, part unspecified: 3
- Malignant neoplasm of prostate: 283
- Malignant neoplasm of upper-outer quadrant of...: 9
- Malignant neoplasm of uterus, part unspecified: 2
- Malignant neoplasm of vagina: 3
- Malignant neoplasm of vulva, unspecified site: 5

Massachusetts Institute of Technology · srlf SOCIÉTÉ DE RÉANIMATION DE LANGUE FRANÇAISE · UNIVERSITÉ PARIS DESCARTES · ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS · www.aphp.fr

Plausibility

- **How patients with pulmonary embolism & without heparine?**
  - Pulmonary embolism : n= 831
  - Septic pulmonary embolism : n= 40
- **How many patients with ARDS without mechanical ventilation?**

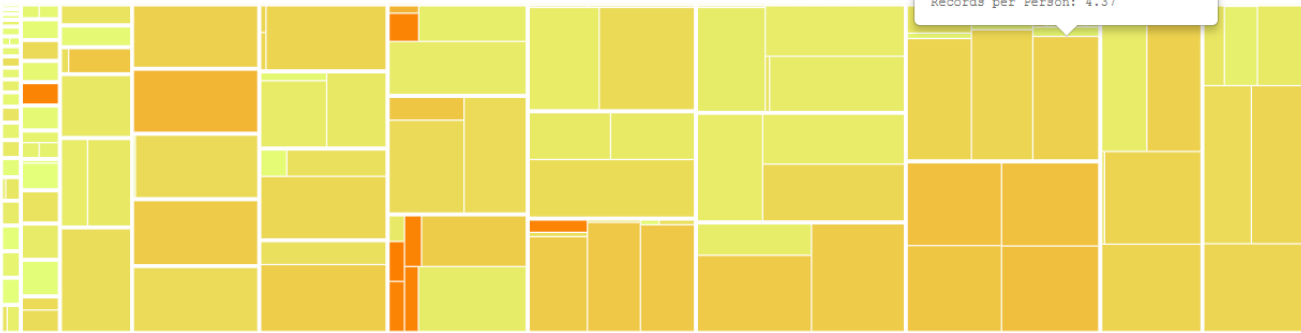- **How many patients with cardiac arrest & without epinephrine/norepinephrine?**
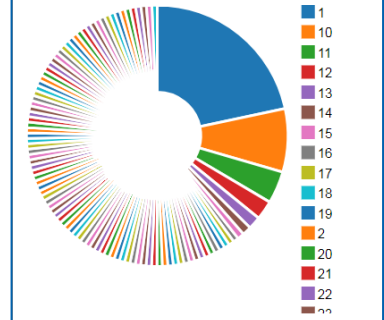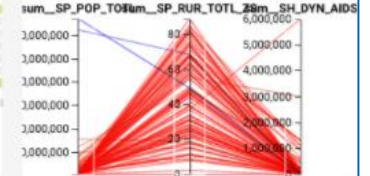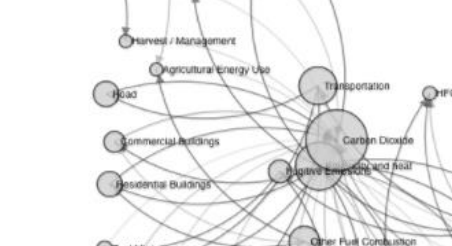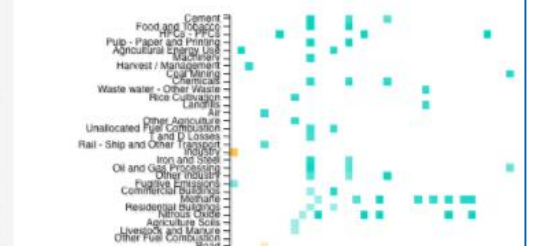
- **How to define the plausibility of the data**

  ▸ "to what extent the data is medically credible"

- **Database certification : comparing data quality scores**

  ▸ **OMOP-based MIMIC** /MIMIC

  ▸ **OMOP-based MIMIC** /other OMOP-based DB of the OHDSI research network

  ▸ **OMOP-based AP-HP DB** /other OMOP-based DB of the OHDSI research network

Massachusetts Institute of Technology

srlf
SOCIÉTÉ
DE RÉANIMATION
DE LANGUE FRANÇAISE

UNIVERSITÉ PARIS DESCARTES

ASSISTANCE PUBLIQUE
HÔPITAUX DE PARIS
www.aphp.fr