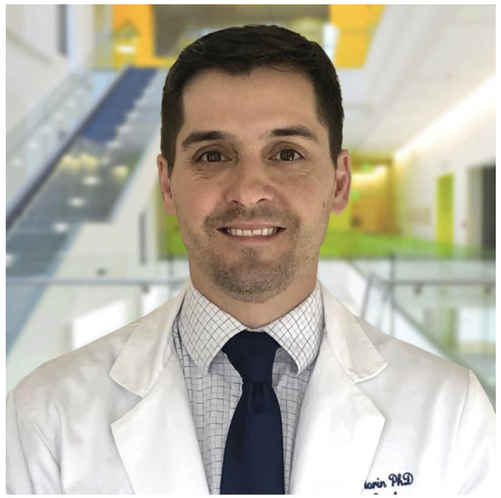
MEDomics

Towards Self-Cognizant Hospitals in the Treatment of Cancer



Olivier Morin, PhD

Chief of Physics, Interim

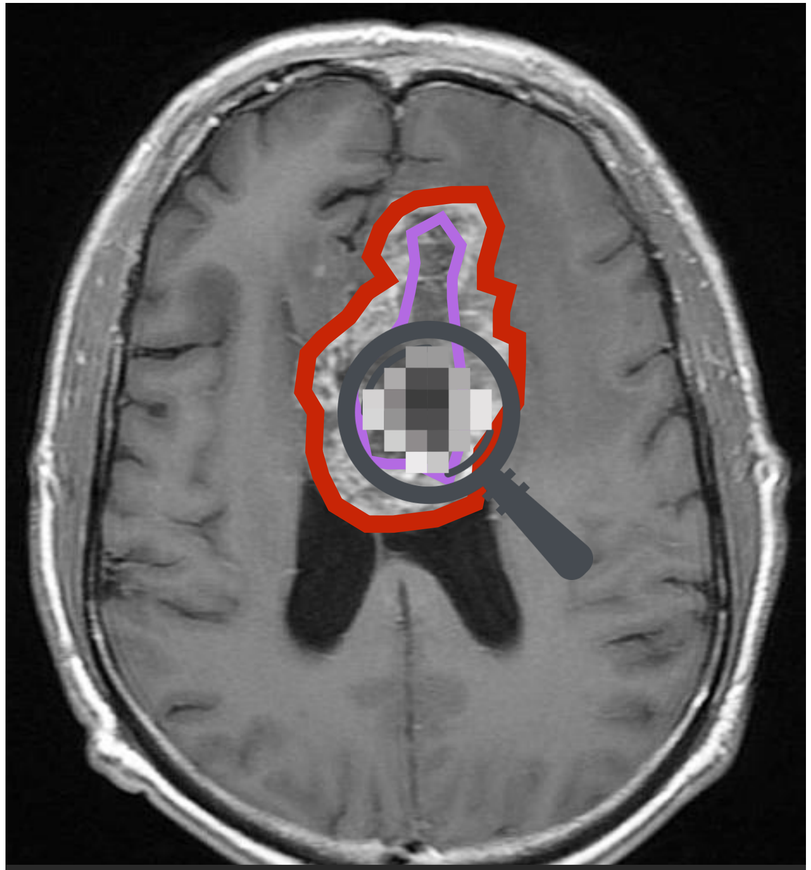
University California San Francisco, CA, USA

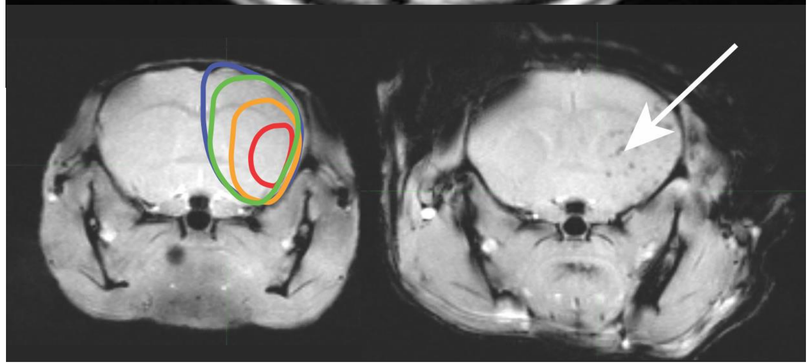
Joint AAPM | COMP Annual Meeting

July 15th, 2020

www.medomics.ai

Lab Efforts





Main Collaborators

UCSF RadOnc:

- **Catherine Park, MD**
- **Jorge Barrios, PhD**
- **Taman Upadhaya, PhD**
- **Steve Braunstein, MD PhD**
- **Sue Yom, MD**
- David Raleigh, MD PhD
- Joe Hsu, MD
- Gilmer Valdes, PhD
- Jean Nakamura, MD
- Jason Chan, MD
- Penny Sneed, MD
- Lijun Ma, PhD
- Benjamin Ziemer, PhD

UCSF Radiology/surgery:

- Javier Villanueva-Meyer MD
- Spencer Behr, MD
- Janine Luppo, PhD
- Antonio Carlos Westphalen, MD
- Michael McDermott, MD

D-Lab/Maastricht U:

- **Philippe Lambin, MD**
- **Simon Keek, PhD student**
- **Henry Woodruff, PhD**
- **Abdalla Ibrahim, MD PhD**
- **Avishek Chatterjee, PhD**

McGill/Sherbrooke:

- **Martin Vallières, PhD**
- **Jan Seuntjens, PhD**

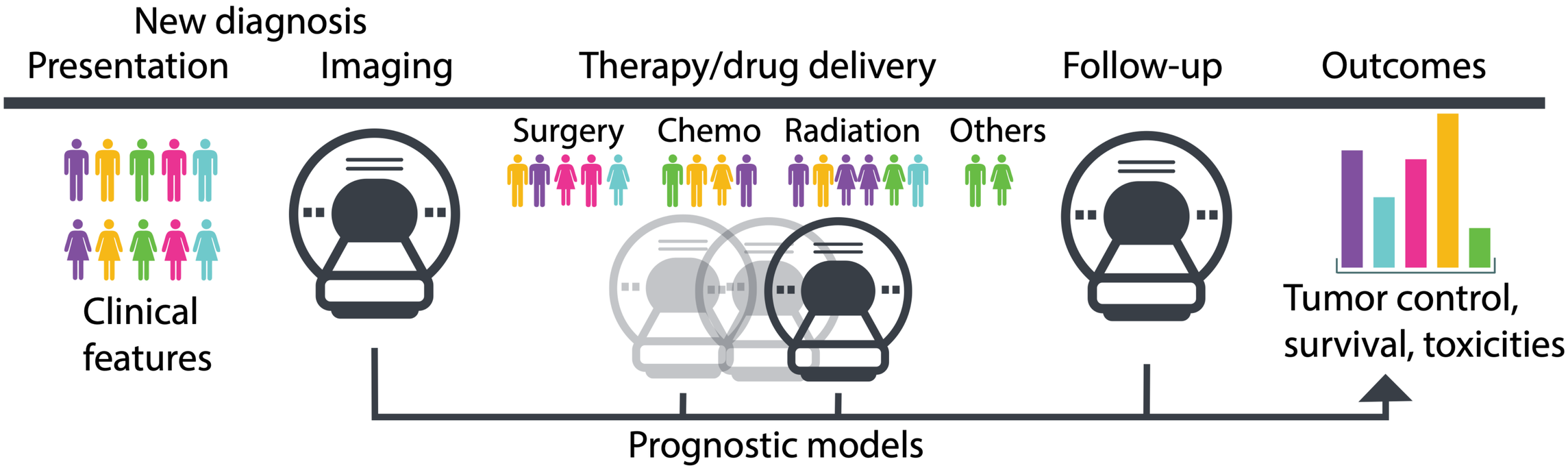
Dresden, Germany

- **Alexander Zwanenburg, PhD**
- **Steffen Lock, PhD**

USF (data science):

- **Yannet Interian, PhD**
- Jeremy Howard, PhD

Point-of-Care Opportunities



- Many statistical models have been developed.
- Few have been integrated in the clinic.
- My lab is studying various point-of-care interventions using informatics.

Oncology

Today

- Increasingly digital, not always accessible, not centralized
- Single shot research
- Transactional
- Advances driven mainly by clinical trials
- Complex, for all
- Frustrating, for patients

Future

- Fully digital and accessible
- Technologies employed for data centralization and governance
- Advances driven by clinical trials but also influenced by real-world data
- **Complexity will be increasingly hidden**
- Quality will be assessed in real-time
- Participatory, data ownership?

Where do we start?

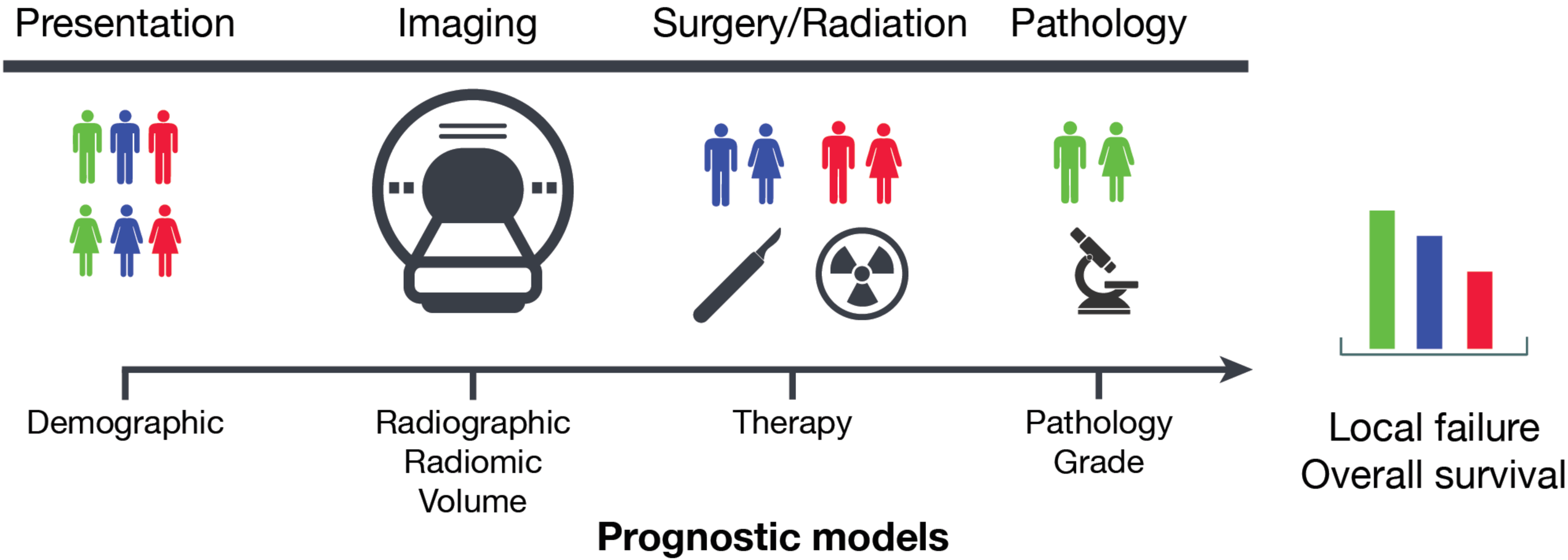
Self-Cognizant Hospitals

having knowledge and being aware of itself and its goals

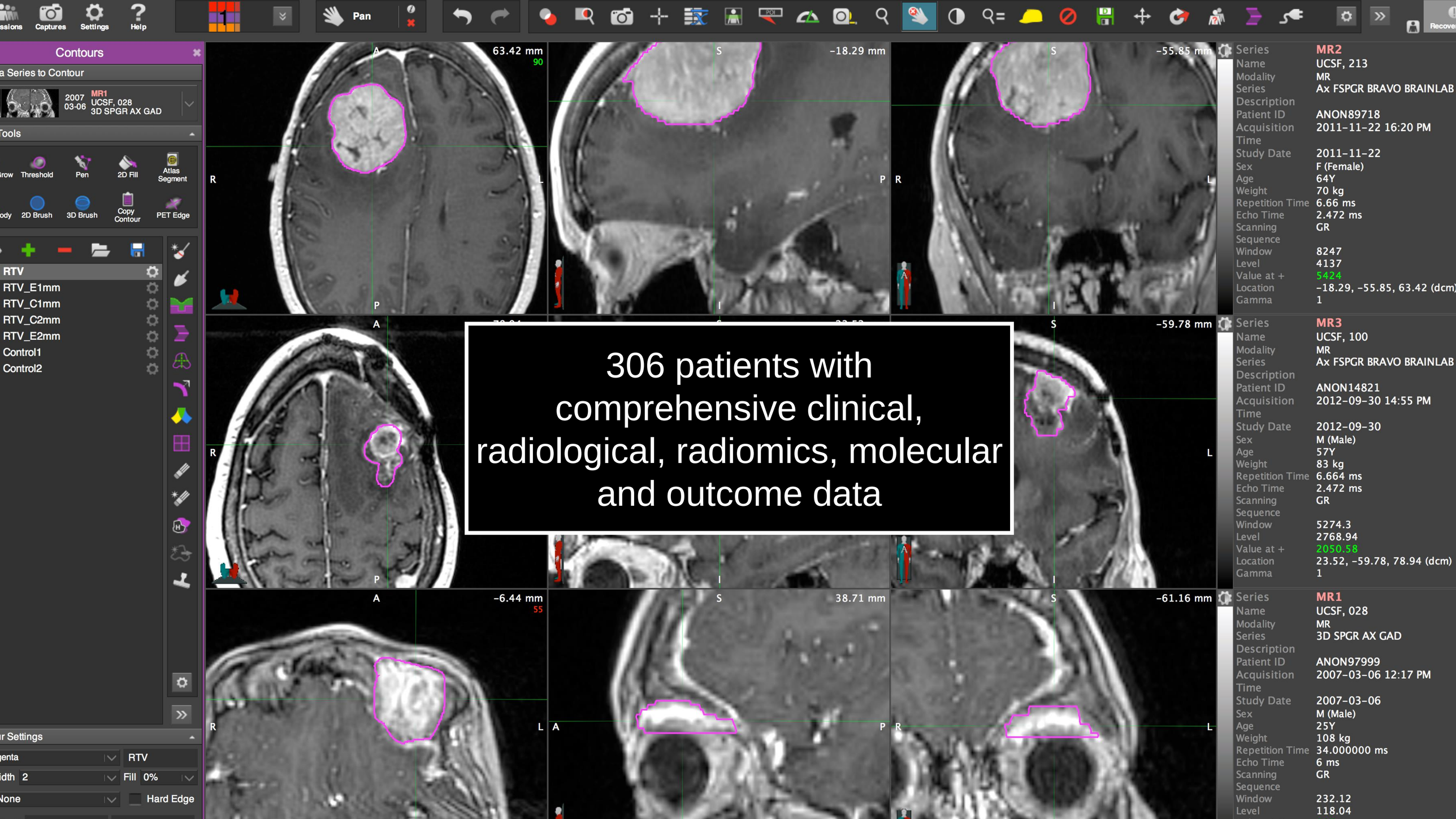
- **Have clear rules on the data needed for each medical intervention and decision.**
- **Missing data will be identified and collected.**
- **Data quality will be assessed and corrected.**
- **Data will be synthesized (MEDomics).**
- **Hospital value/cost and performance (patient quality of life) will be measured and compared to regional/national/international trends.**

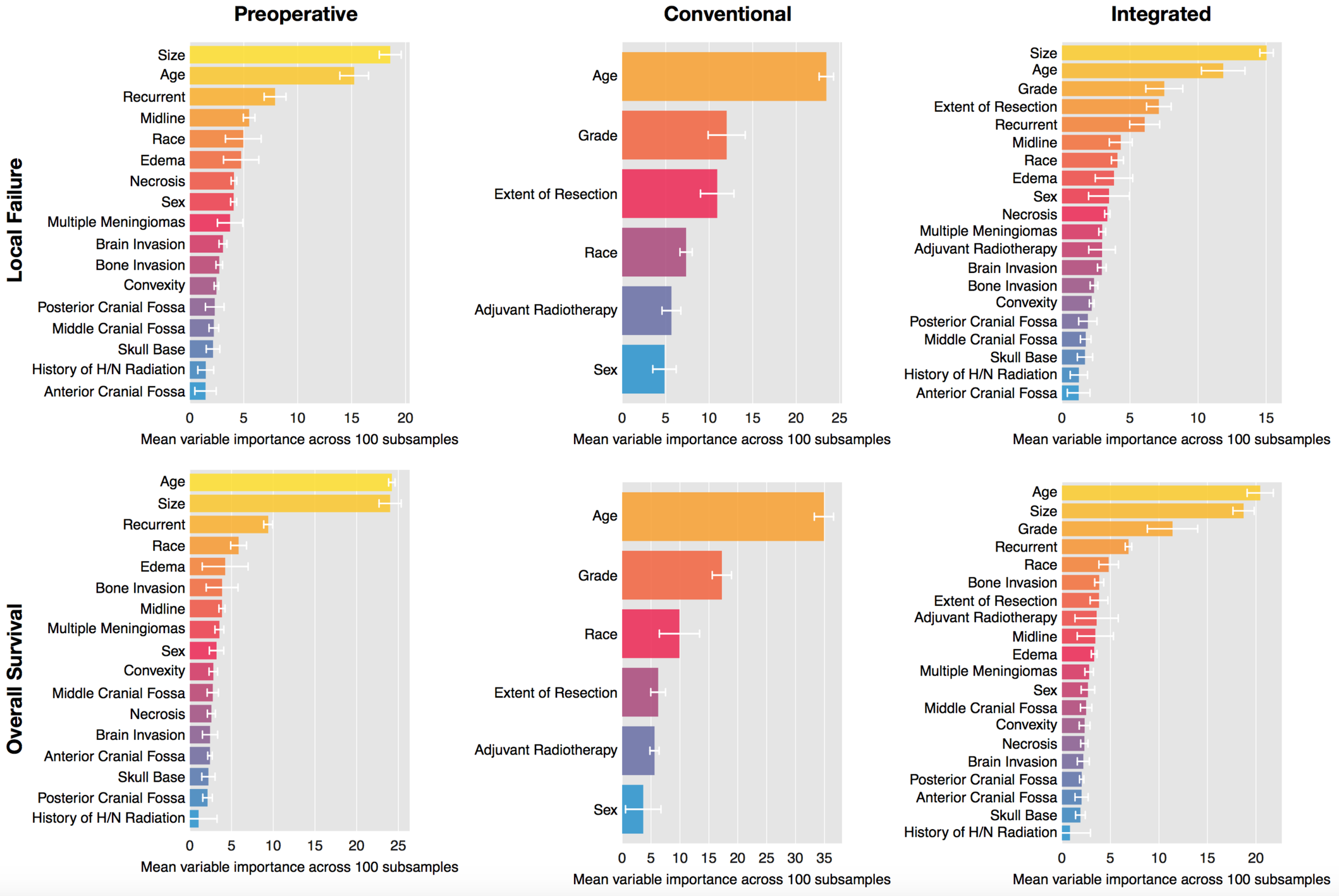
MEDomics Animation: https://youtu.be/2030Pdgm3_4

Meningioma example



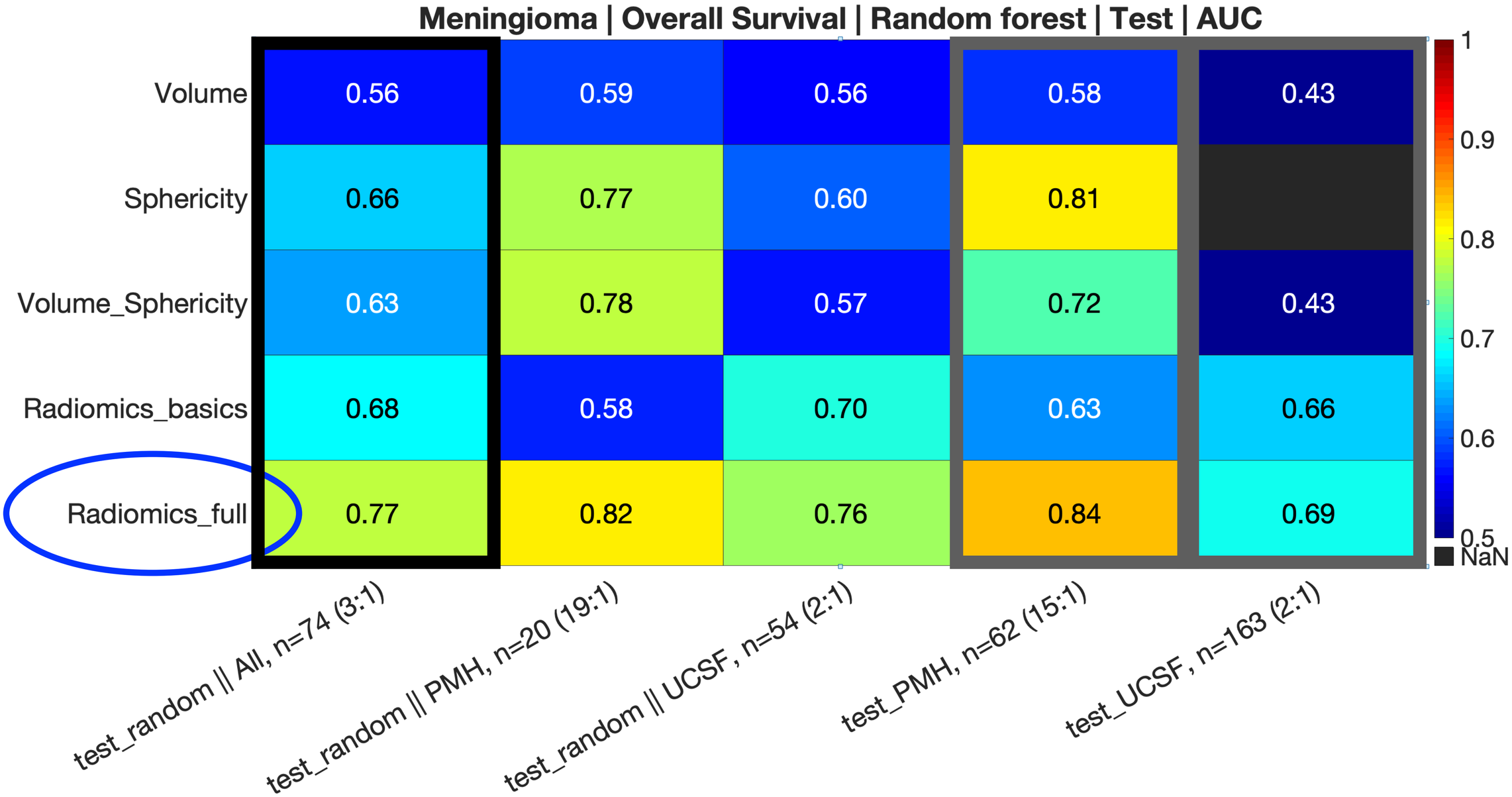
Morin O, Chen WC, Nassiri F, et al. Integrated models incorporating radiologic and radiomic features predict meningioma grade, local failure, and overall survival. *Neurooncol Adv* 2019; 1(1): vdz011.

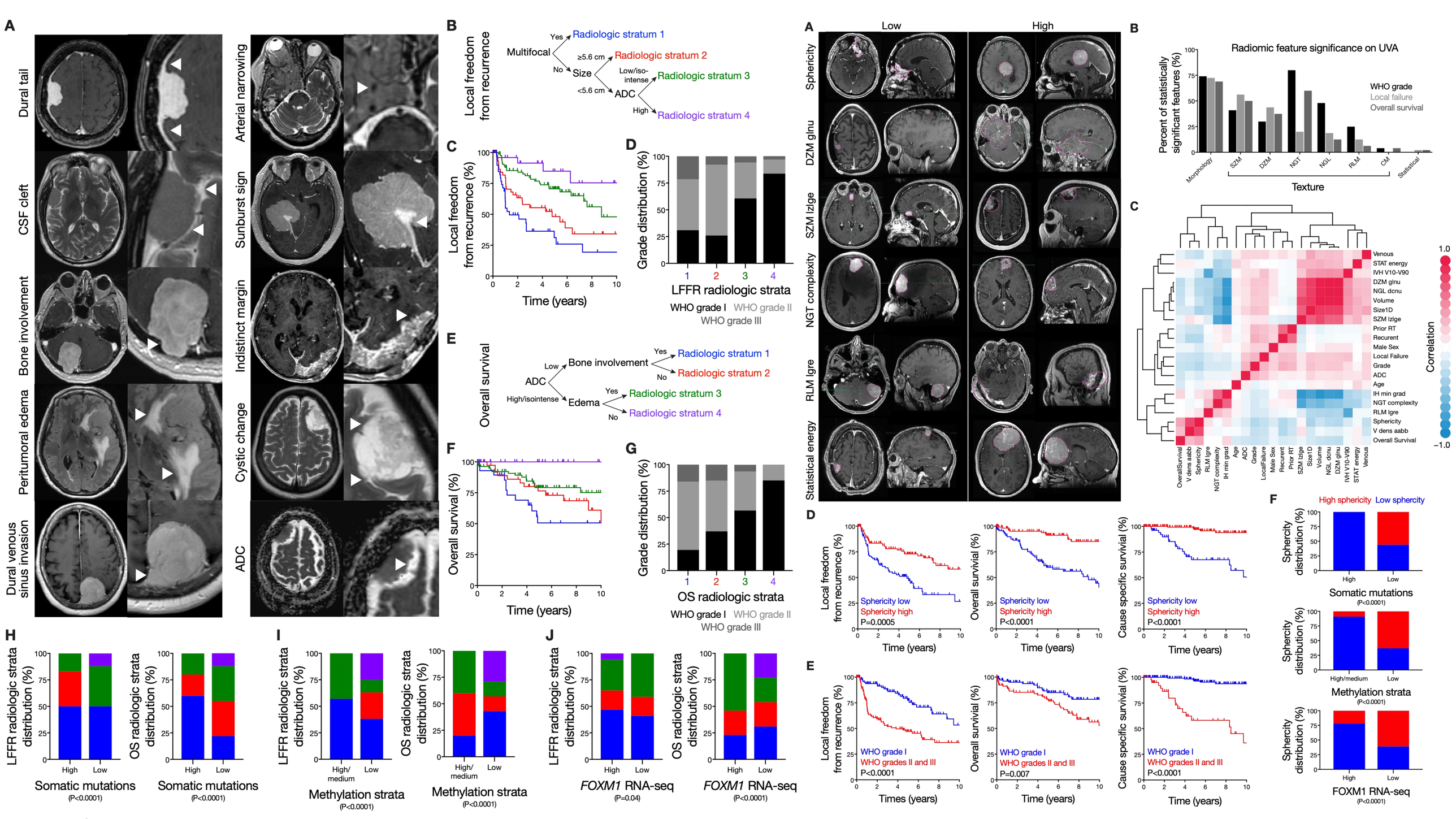

Contours
Series to Contour
2007 03-06 MR1 UCSF, 028 3D SPGR AX GAD
Tools
Threshold
Pen
2D Fill
Atlas Segment
2D Brush
3D Brush
Copy Contour
PET Edge
RTV
RTV_E1mm
RTV_C1mm
RTV_C2mm
RTV_E2mm
Control1
Control2
Settings
RTV
Fill 0%
Hard Edge
63.42 mm
-18.29 mm
-55.85 mm
-59.78 mm
-6.44 mm
38.71 mm
-61.16 mm
306 patients with comprehensive clinical, radiological, radiomics, molecular and outcome data
Series MR2
Name UCSF, 213
Modality MR
Series Ax FSPGR BRAVO BRAINLAB
Description
Patient ID ANON89718
Acquisition Time 2011-11-22 16:20 PM
Study Date 2011-11-22
Sex F (Female)
Age 64Y
Weight 70 kg
Repetition Time 6.66 ms
Echo Time 2.472 ms
Scanning Sequence GR
Window 8247
Level 4137
Value at + 5424
Location -18.29, -55.85, 63.42 (dcm)
Gamma 1
Series MR3
Name UCSF, 100
Modality MR
Series Ax FSPGR BRAVO BRAINLAB
Description
Patient ID ANON14821
Acquisition Time 2012-09-30 14:55 PM
Study Date 2012-09-30
Sex M (Male)
Age 57Y
Weight 83 kg
Repetition Time 6.664 ms
Echo Time 2.472 ms
Scanning Sequence GR
Window 5274.3
Level 2768.94
Value at + 2050.58
Location 23.52, -59.78, 78.94 (dcm)
Gamma 1
Series MR1
Name UCSF, 028
Modality MR
Series 3D SPGR AX GAD
Description
Patient ID ANON97999
Acquisition Time 2007-03-06 12:17 PM
Study Date 2007-03-06
Sex M (Male)
Age 25Y
Weight 108 kg
Repetition Time 34.000000 ms
Echo Time 6 ms
Scanning Sequence GR
Window 232.12
Level 118.04


Preoperative
Conventional
Integrated
Local Failure
Overall Survival
Size
Age
Recurrent
Midline
Race
Edema
Necrosis
Sex
Multiple Meningiomas
Brain Invasion
Bone Invasion
Convexity
Posterior Cranial Fossa
Middle Cranial Fossa
Skull Base
History of H/N Radiation
Anterior Cranial Fossa
Grade
Extent of Resection
Adjuvant Radiotherapy
0 5 10 15 20
0 5 10 15 20 25
0 5 10 15
0 5 10 15 20 25
0 5 10 15 20 25 30 35
Mean variable importance across 100 subsamples

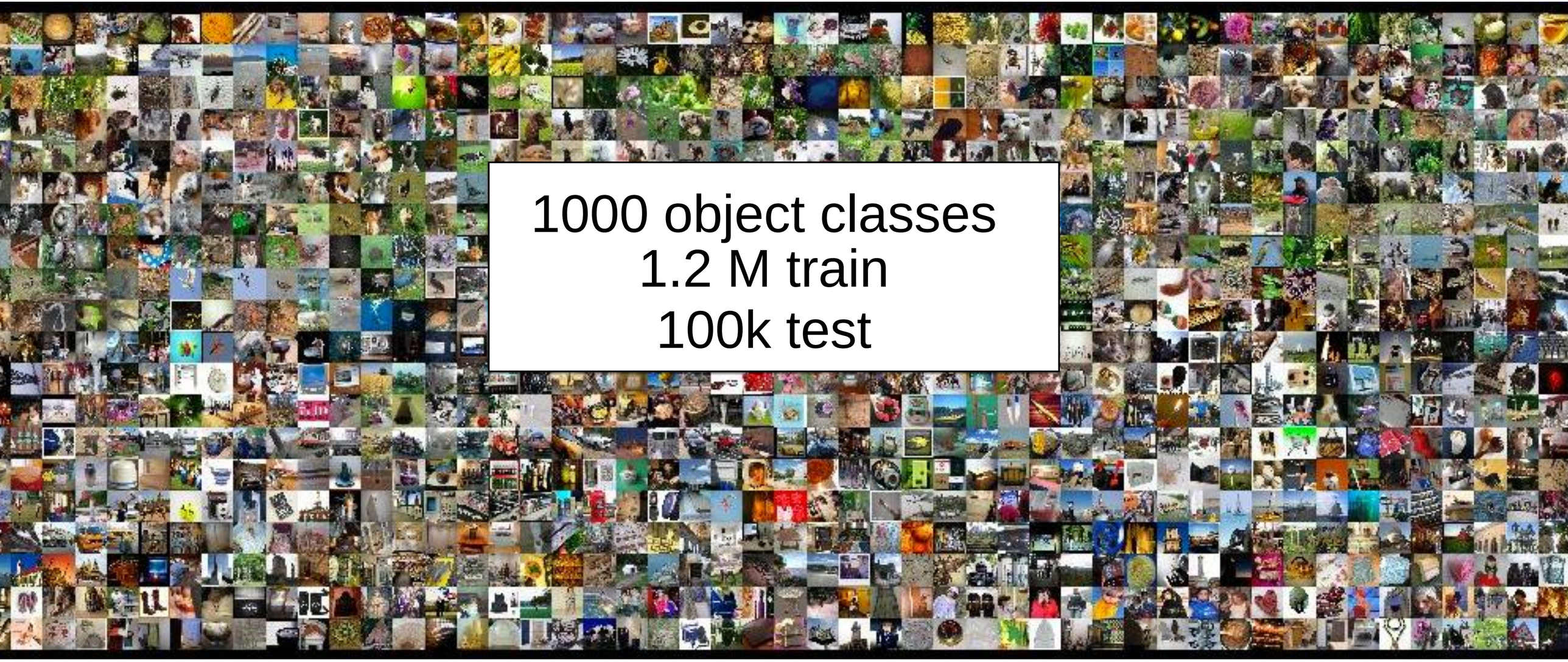
RADIOMICS BASELINE

Model analysis module

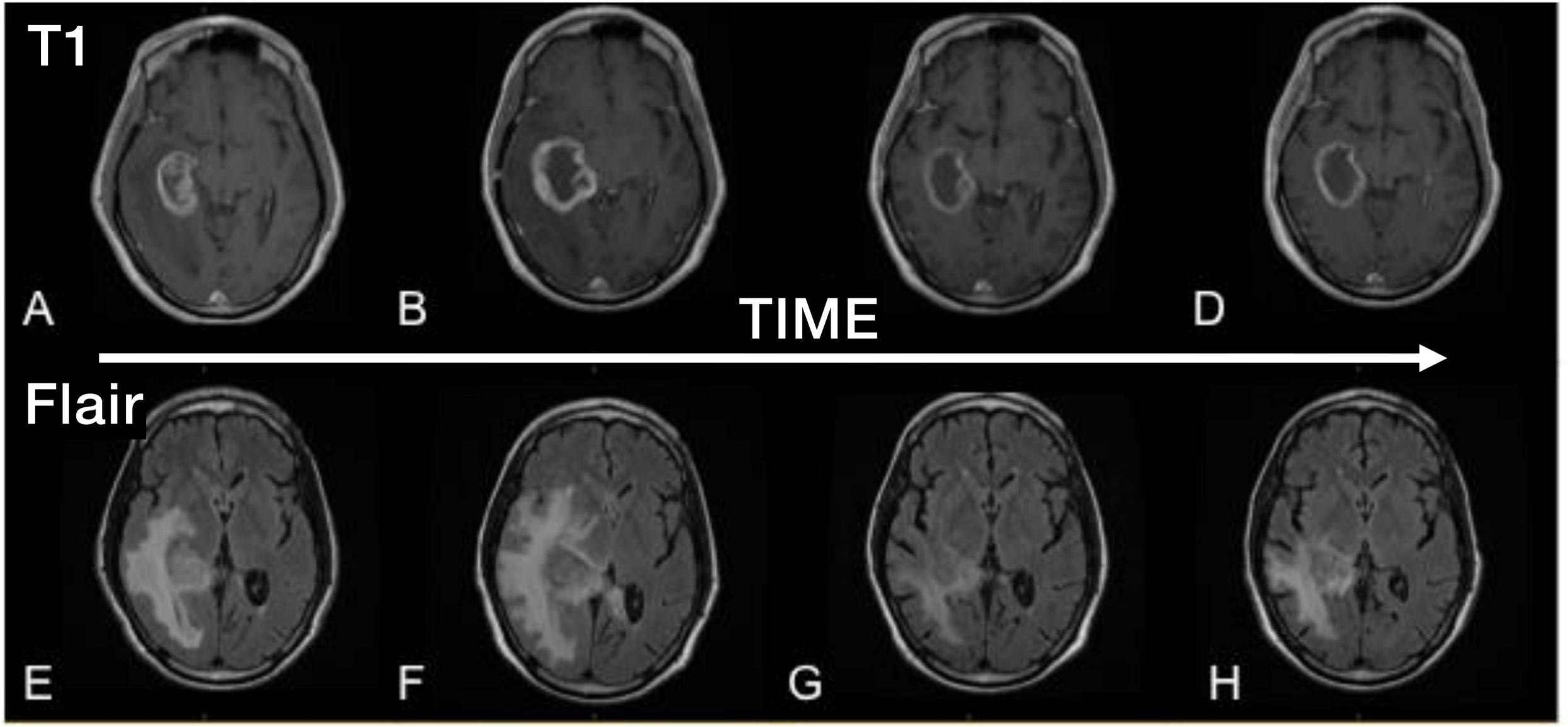



A
Dural tail
CSF cleft
Bone involvement
Peritumoral edema
Dural venous sinus invasion
Arterial narrowing
Sunburst sign
Indistinct margin
Cystic change
ADC
B
Local freedom from recurrence
Multifocal
Yes
Radiologic stratum 1
No
Size
≥5.6 cm
Radiologic stratum 2
<5.6 cm
ADC
Low/iso-intense
Radiologic stratum 3
High
Radiologic stratum 4
C
Local freedom from recurrence (%)
Time (years)
D
Grade distribution (%)
LFFR radiologic strata
WHO grade I
WHO grade II
WHO grade III
E
Overall survival
ADC
Low
Bone involvement
Yes
Radiologic stratum 1
No
Radiologic stratum 2
High/isointense
Edema
Yes
Radiologic stratum 3
No
Radiologic stratum 4
F
Overall survival (%)
Time (years)
G
Grade distribution (%)
OS radiologic strata
WHO grade I
WHO grade II
WHO grade III
H
LFFR radiologic strata distribution (%)
High
Low
Somatic mutations
(P<0.0001)
OS radiologic strata distribution (%)
High
Low
Somatic mutations
(P<0.0001)
I
LFFR radiologic strata distribution (%)
High/ medium
Low
Methylation strata
(P<0.0001)
OS radiologic strata distribution (%)
High/ medium
Low
Methylation strata
(P<0.0001)
J
LFFR radiologic strata distribution (%)
High
Low
FOXM1 RNA-seq
(P=0.04)
OS radiologic strata distribution (%)
High
Low
FOXM1 RNA-seq
(P<0.0001)
A
Low
High
Sphericity
DZM glnu
SZM lzlge
NGT complexity
RLM lgre
Statistical energy
B
Radiomic feature significance on UVA
Percent of statistically significant features (%)
WHO grade
Local failure
Overall survival
Morphology
SZM
DZM
NGT
NGL
RLM
CM
Statistical
Texture
C
Venous
STAT energy
IVH V10-V90
DZM glnu
NGL dcnu
Volume
Size1D
SZM lzlge
Prior RT
Recurent
Male Sex
Local Failure
Grade
ADC
Age
IH min grad
NGT complexity
RLM lgre
Sphericity
V dens aabb
Overall Survival
Correlation
1.0
−1.0
D
Local freedom from recurrence (%)
Sphericity low
Sphericity high
P=0.0005
Time (years)
Overall survivial (%)
Sphericity low
Sphericity high
P<0.0001
Time (years)
Cause specific survivial (%)
Sphericity low
Sphericity high
P<0.0001
Time (years)
E
Local freedom from recurrence (%)
WHO grade I
WHO grades II and III
P<0.0001
Times (years)
Overall survivial (%)
WHO grade I
WHO grades II and III
P=0.007
Time (years)
Cause specific survivial (%)
WHO grade I
WHO grades II and III
P<0.0001
Time (years)
F
High sphericity
Low sphericity
Sphericity distribution (%)
High
Low
Somatic mutations
(P<0.0001)
Sphericity distribution (%)
High/medium
Low
Methylation strata
(P<0.0001)
Sphericity distribution (%)
High
Low
FOXM1 RNA-seq
(P<0.0001)

We need data!



We need diverse data with labels!



Data Flow

Morin et al, A Deep Look into the Future of Quantitative Imaging in Oncology: A Statement of Working Principles and Proposal for Change, Special Imaging Issue, International Journal of Radiation Oncology, Biology, Physics, 2018.

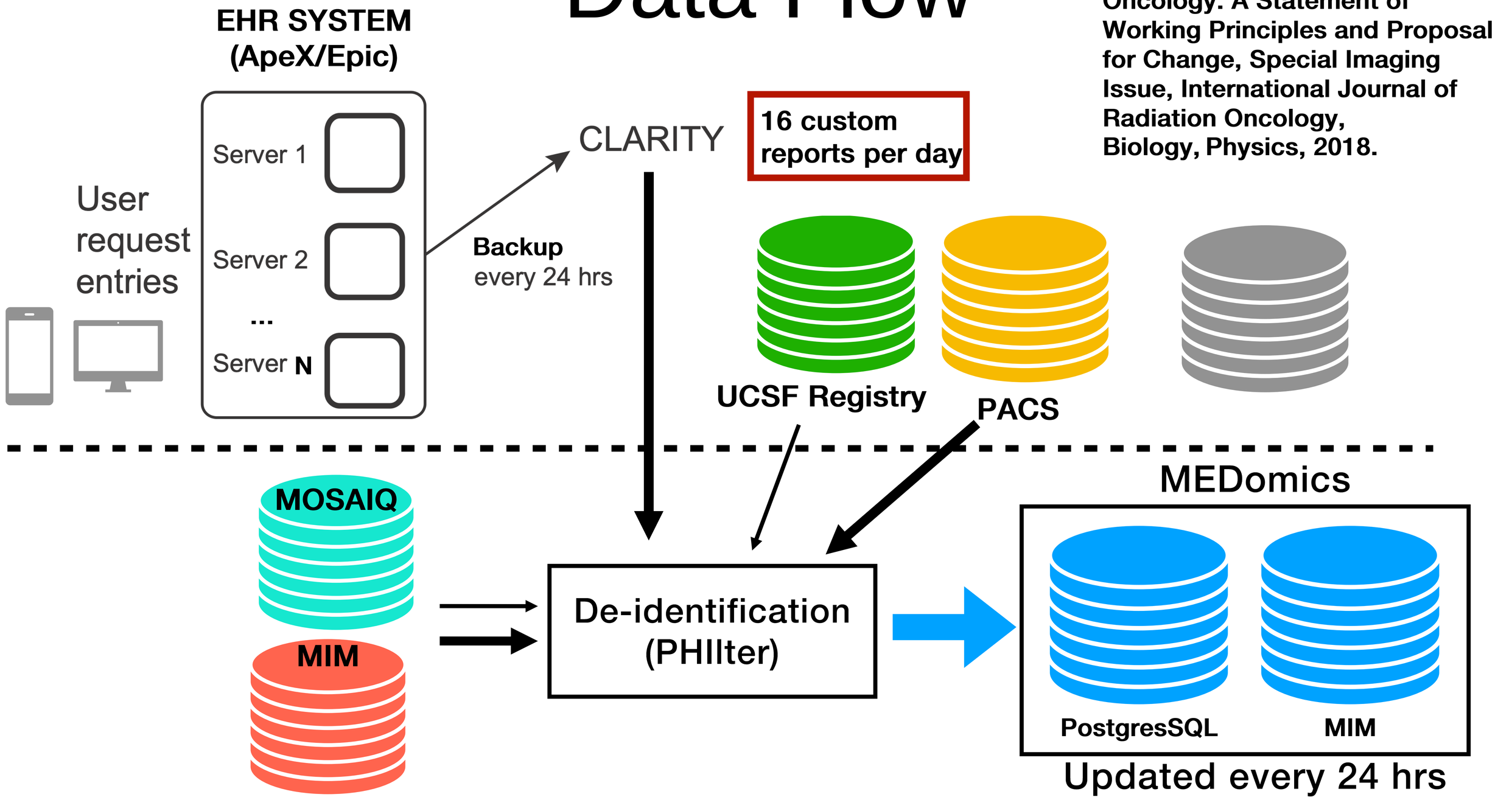


TABLE	# ROWS
Allergy	150k
Encounters	3M
Family	430k
Imaging	1.7M
Labs	200M
Medical	670k
Medication	15M
Micro	9M
Notes	12.2M
Pathology	860k
Patient	159k
Problem	681k
Procedures	89k
Social	115k
Surgical	300k
Radiation	42k
Charges	636M

IRB Number: 18-25441, 18-26489

ount ▼	ICD10CM	Description
	I10	Essential (primary) hypertension
	C50919	Malignant neoplasm of unspecified site of unspecified female breast
	C61	Malignant neoplasm of prostate
	Z7189	Other specified counseling
	C73	Malignant neoplasm of thyroid gland
	Z0000	Encounter for general adult medical examination without abnormal findings
	C719	Malignant neoplasm of brain, unspecified
	E785	Hyperlipidemia, unspecified
	C3490	Malignant neoplasm of unspecified part of unspecified bronchus or lung
	R52	Pain, unspecified
	K5900	Constipation, unspecified
	Z66	Do not resuscitate
	D649	Anemia, unspecified
	E119	Type 2 diabetes mellitus without complications
	C439	Malignant melanoma of skin, unspecified
	D496	Neoplasm of unspecified behavior of brain
	Z5111	Encounter for antineoplastic chemotherapy
	F329	Major depressive disorder, single episode, unspecified
	E039	Hypothyroidism, unspecified
	K219	Gastro-esophageal reflux disease without esophagitis
	Z789	Other specified health status
612	F419	Anxiety disorder, unspecified

Queries ⌘1 | Schema ⌘2 | Find immunotherapy drug hits

Filter schema... qip...

Tables
- allergy
- demographic
- encounters
- familyhistory
- imaging
- labs
- medicalhistory
- medication
- micro
- notes
- pathology
- problemlist
- procedure
- socialhistory
- surgicalhistory

Views

x_2 Functions

Find immunotherapy drug hits

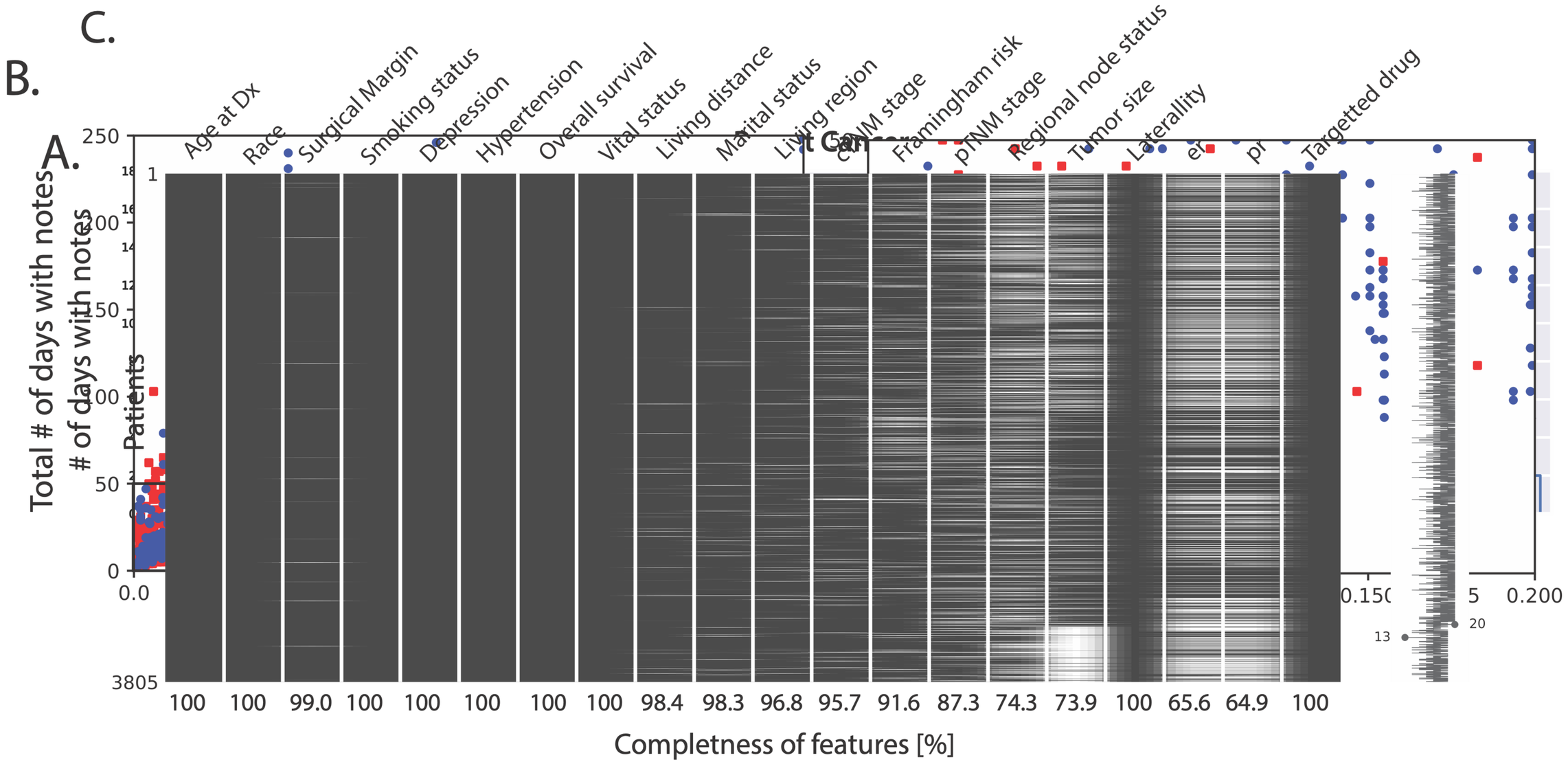
```
1  select demographic.mrn, medication_startdate, medication_name, medication_route from qipm.demographic
2  join qipm.medication on demographic.mrn = medication.mrn
3  where medication_name ilike '%Ipilimumab%'
4  or medication_name ilike '%Nivolumab%'
5  or medication_name ilike '%Pembrolizumab%'
6  or medication_name ilike '%Atezolizumab%'
7  or medication_name ilike '%Avelumab%'
8  or medication_name ilike '%Durvalumab%'
```

Success
2799 rows

Explore | SQL | Data | Chart | Export

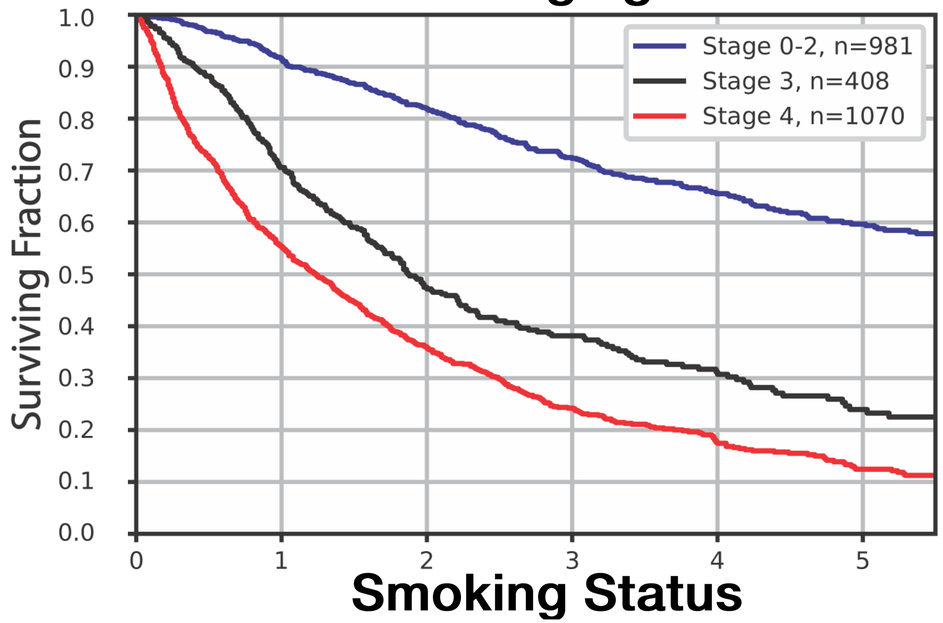
mrn	medication_startdate	medication_name	medication_route
	2018-03-14	NIVOLUMAB 100 ML IVPB	Intravenous
	2018-02-21	PEMBROLIZUMAB 50 ML IVPB	Intravenous
	2017-12-27	PEMBROLIZUMAB 50 ML IVPB	Intravenous
	2017-12-27	NIVOLUMAB 100 ML IVPB	Intravenous
	2017-09-27	NIVOLUMAB 100 ML IVPB	Intravenous
	2018-01-26	NIVOLUMAB 100 ML IVPB	Intravenous
	2017-09-19	PEMBROLIZUMAB 50 ML IVPB	Intravenous

Careful Patient Selection



Lung Cancer

Staging



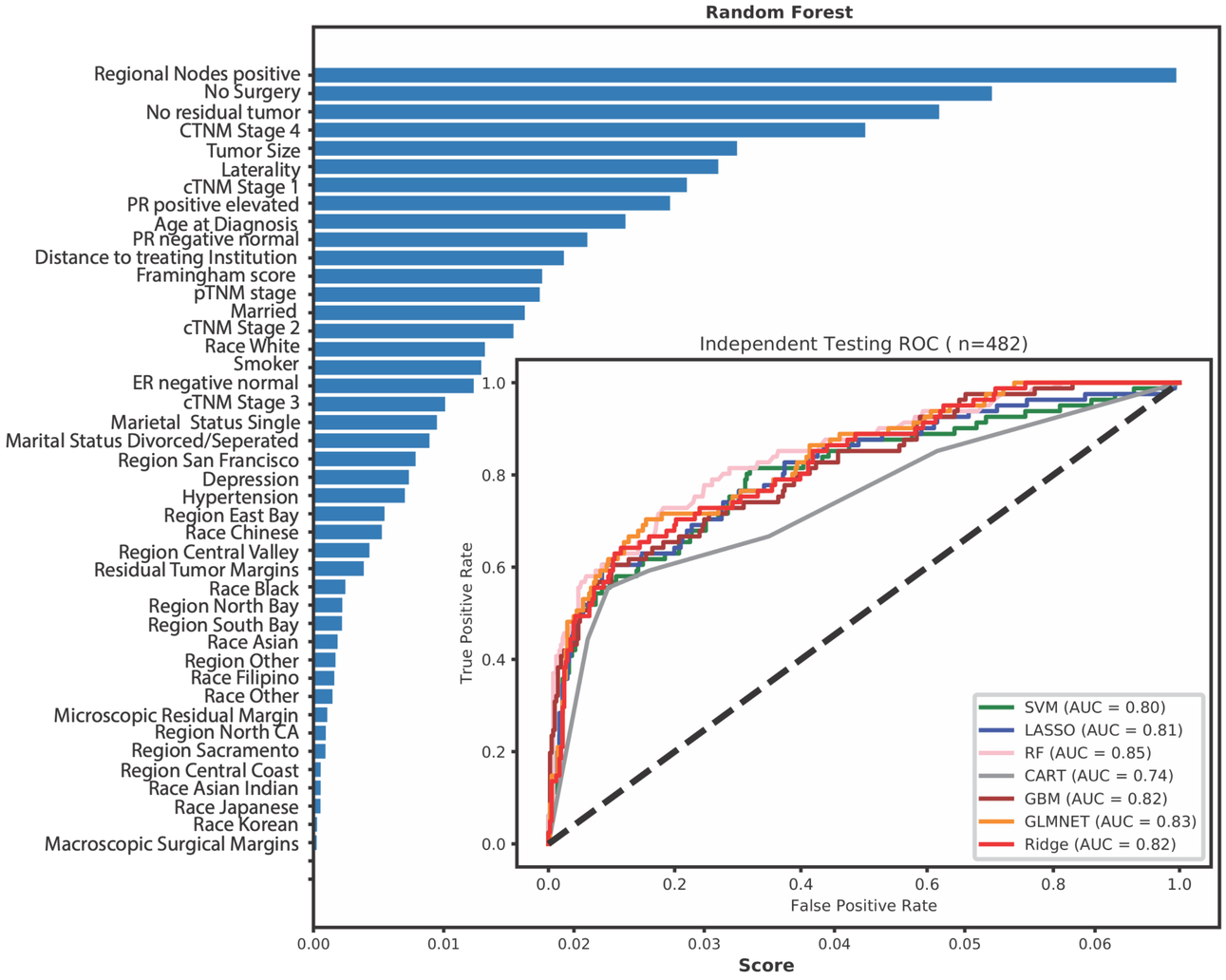
Immunotherapy

Smoking Status

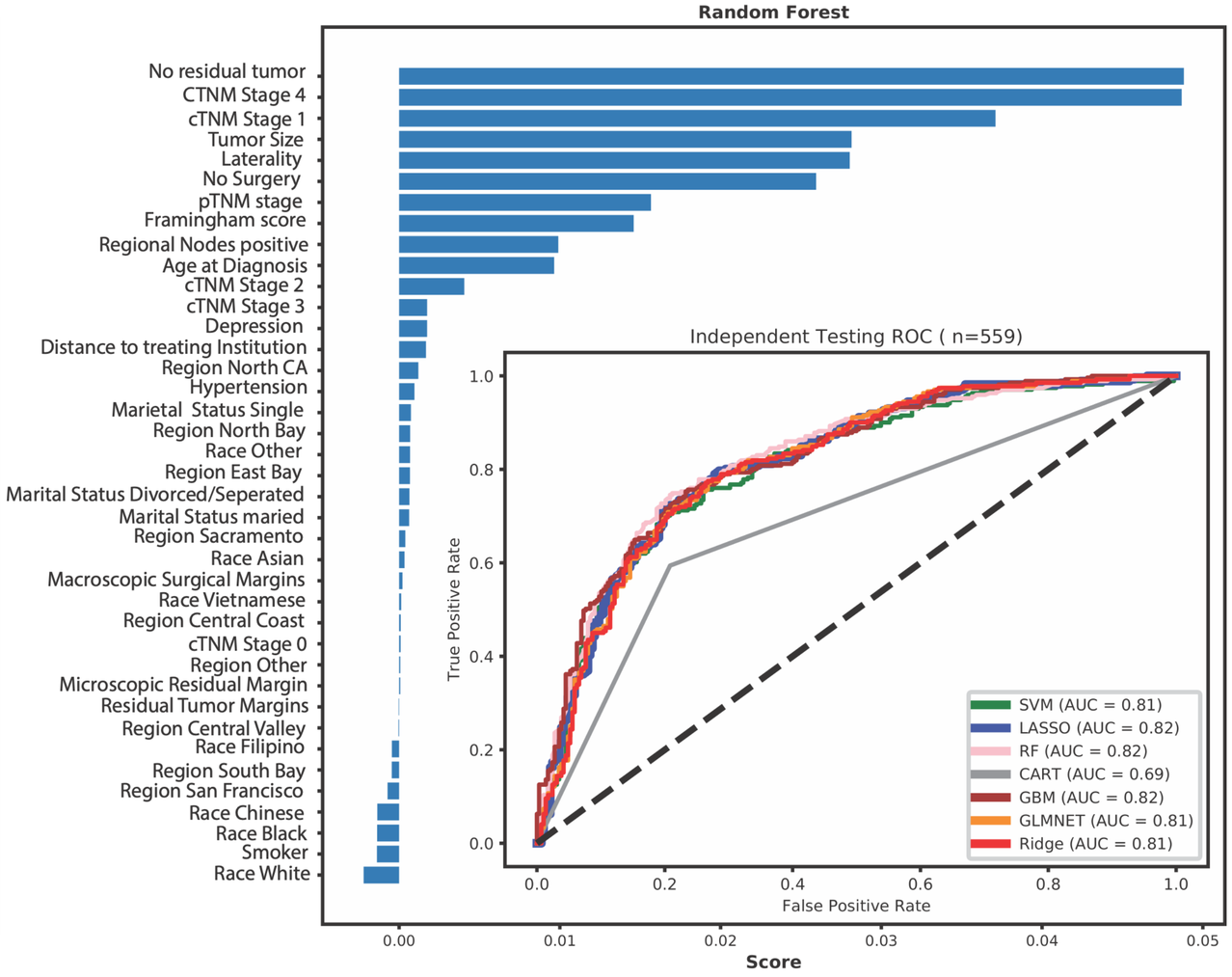
Framingham Risk

MEDomics: preliminary results

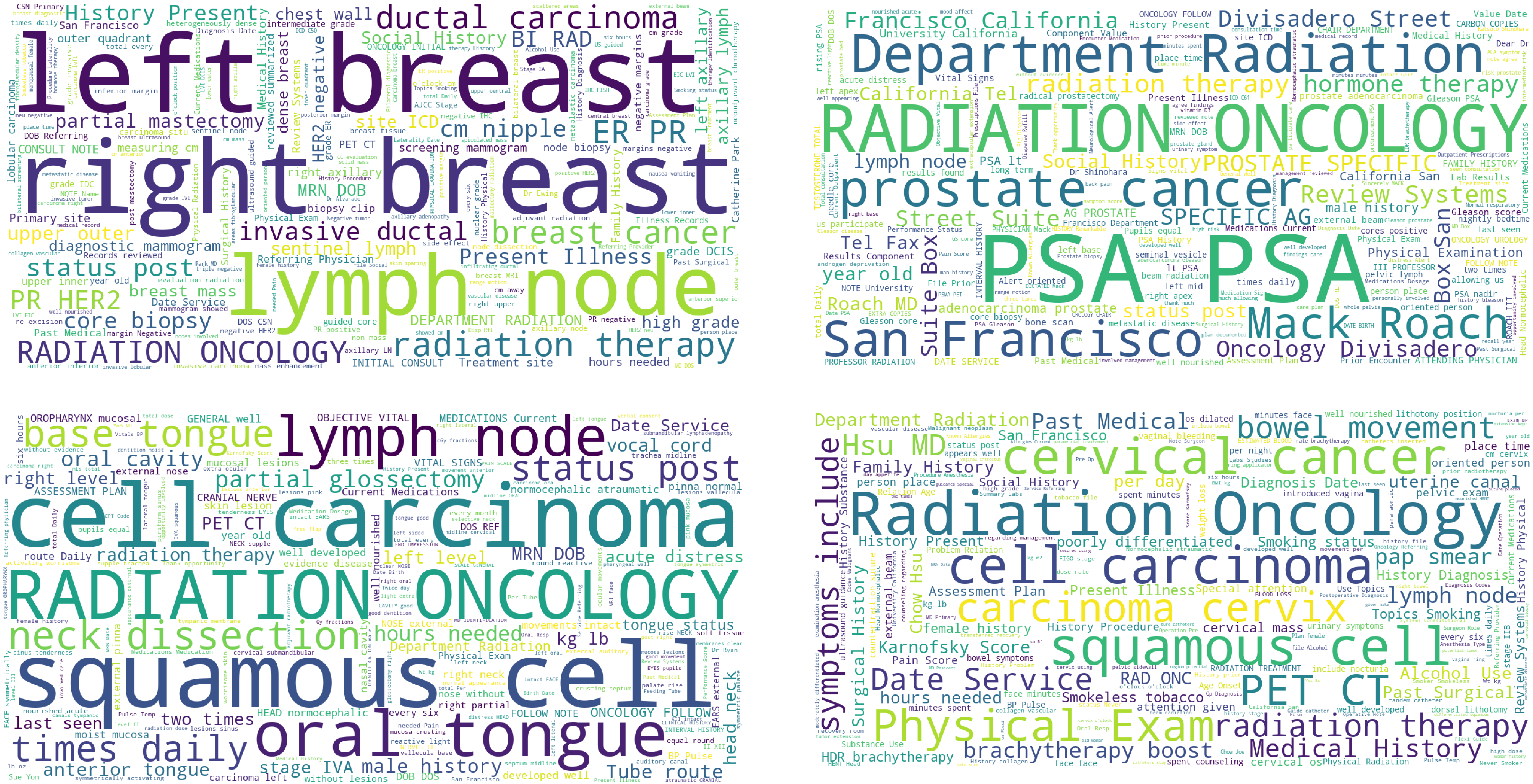
Breast Cancer



Lung Cancer

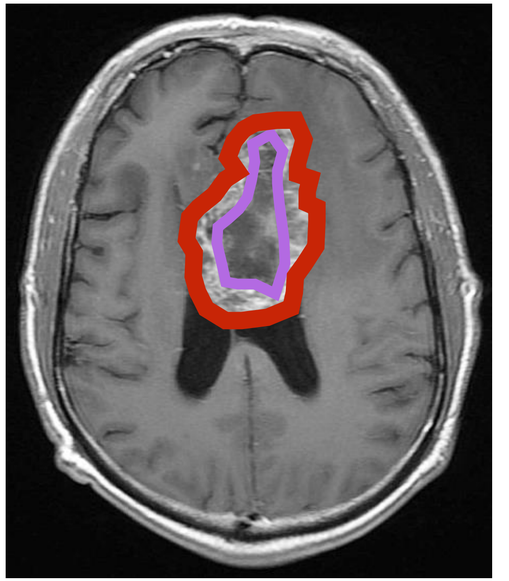


Exploring Medical Notes



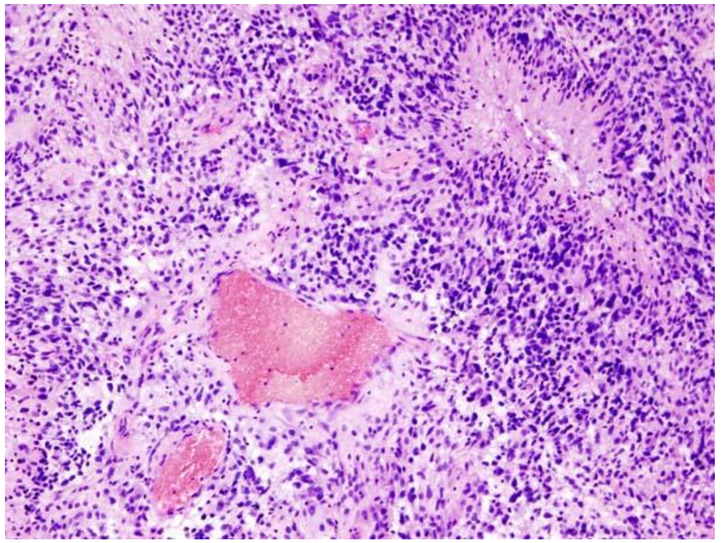
Example of radiology extract:

CT BRAIN WITHOUT CONTRAST: 4/23/2018 2:57 AM CLINICAL HISTORY: Concern for change in ventricular size. COMPARISON: CT brain 4/22/2018. TECHNIQUE: Helical CT imaging of the brain without intravenous contrast. Coronal and sagittal reformatted images were obtained. CONTRAST MEDIA: None. RADIATION DOSE INDICATORS: 1 exposure event(s), CTDIvol: 54.2 mGy. DLP: 1058 mGy-cm. FINDINGS: Parenchyma: No new intracranial hemorrhage or large vessel territory hypodensity. Left frontal EVD with tip terminating near the foramen of Monro, unchanged. Sequela of prior aneurysm clipping. Unchanged appearance of the sunken right cerebral hemisphere. Ventricles: Unchanged lateral ventricular size is Extra-axial collection: Unchanged size and appearance of left-sided extra-axial collection with previous SEPS device. Orbits and globes: Unremarkable. Paranasal sinuses and mastoid air cells: Fluid in the bilateral mastoid air cells. Bones: Sequela of prior right hemicraniectomy. Soft tissues: Unremarkable



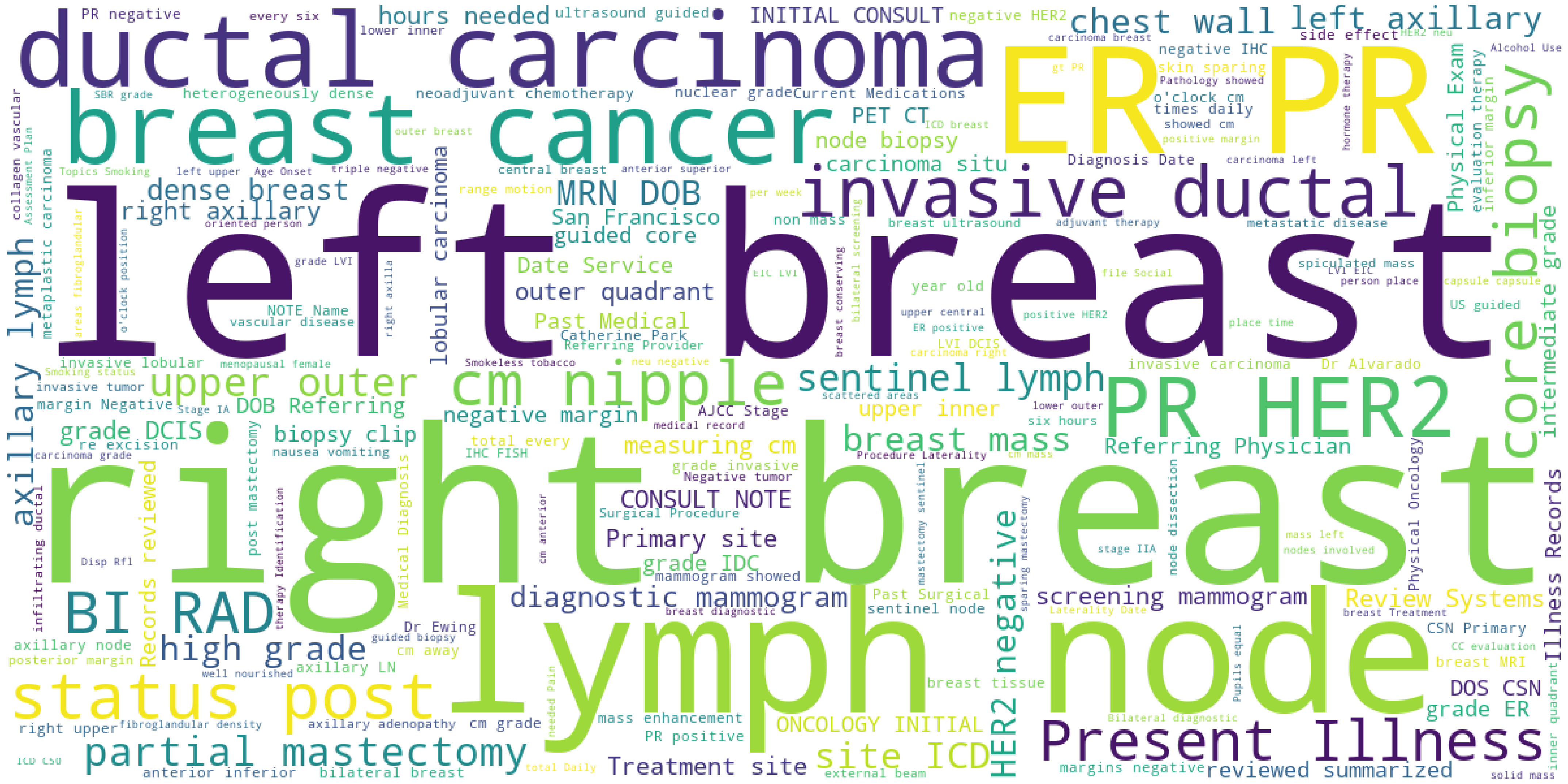
Example of pathology extract:

UNIVERSITY OF CALIFORNIA SAN FRANCISCO DEPARTMENT OF PATHOLOGY 1600 DIVISADERO STREET Room: R-200, Box: 1785 SAN FRANCISCO, CA 94115-1785 TEL: (415) 885-7301 FAX: (415) 353-7676 CYTOPATHOLOGY REPORT Patient Name:XXXXXXXXXXX J. Med. Rec.#: XXXXXX DOB: XXXXXX (Age: 32) Sex: X Accession #: XXXXXX Visit #: XXXXXX Service Date: 4/18/2018 Received: 4/19/2018 Location: SU8 Client: Parnassus Physician(s): XXXXXXXXXX ((415) 885-7788) Source of Specimen: Cervical, ThinPrep FINAL CYTOLOGIC INTERPRETATION/RESULT: A: Cervical, ThinPrep ATYPICAL SQUAMOUS CELLS OF UNDETERMINED SIGNIFICANCE (ASC-US).   SPECIMEN ADEQUACY: Satisfactory for evaluation. Transformation zone components are present. COMMENTS: The specimen will be sent for high risk HPV testing with HPV16/18 genotyping. Please refer to the separate report for results. Clinical History Date of Last Menstrual Period: 4/7/2018 1 Thin Prep Pap. Co-test with reflex 16/18 genotyping in women 30 years and older, if Pap negative and high risk HPV is positive History Abnormal Pap?:No Treatment:Not Applicable Submitting Diagnosis: ICD-10-CM:Z12.4:Encounter for screening for malignant neoplasm of cervix Number of slides: 1 The Pap test is a screening test to aid in the detection of anogenital cancers and their precursors. Both false-negative and false-positive results have been experienced. The Pap should not be used as the sole means to detect anogenital cancers; regular periodic testing and follow-up of unexplained clinical signs and symptoms is suggested.   XXXXXXXX/Pathologist Electronically signed out on 4/23/2018 17:2

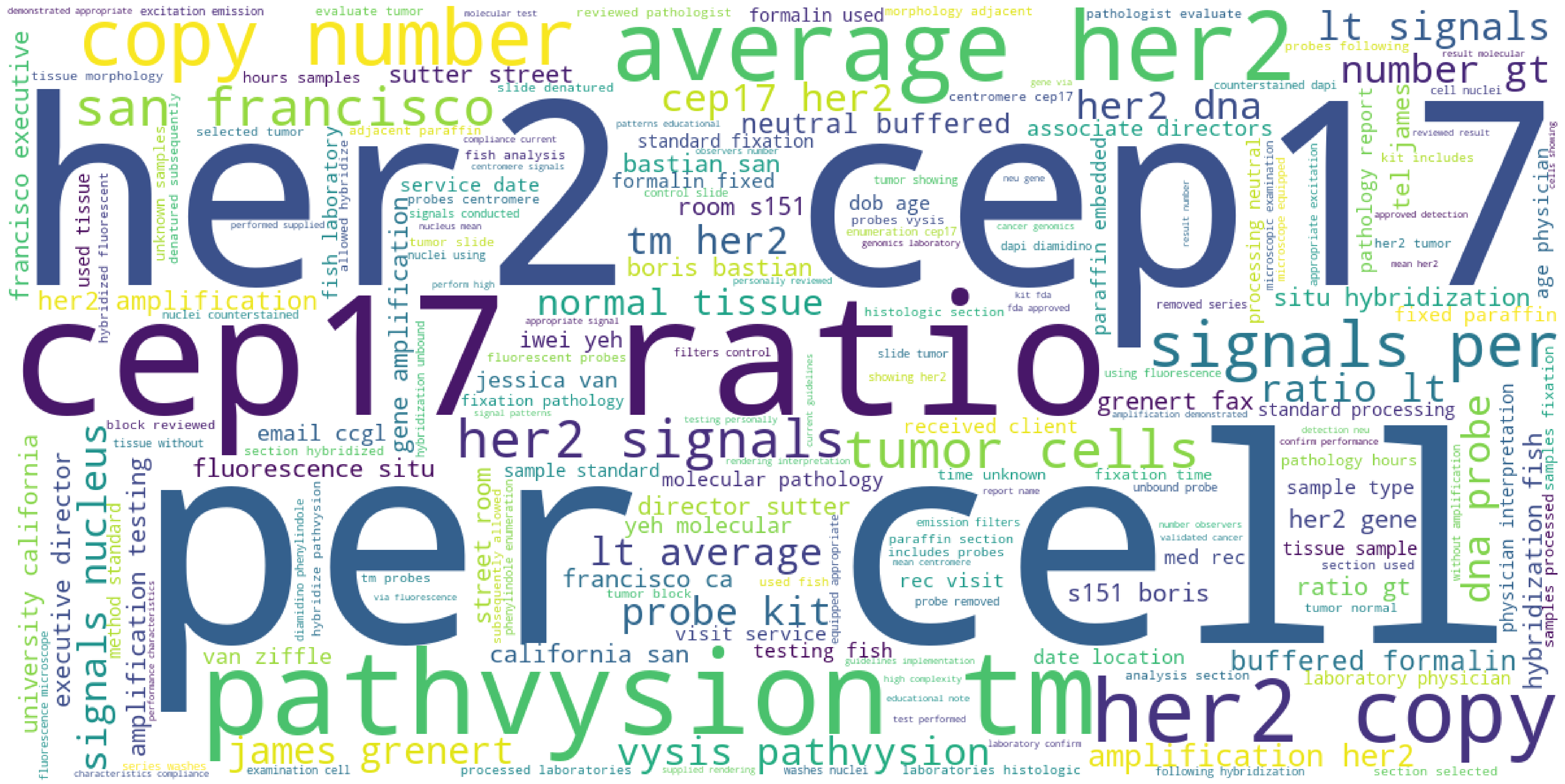


Need Feature Extraction -> Pattern Recognition -> Deep learning

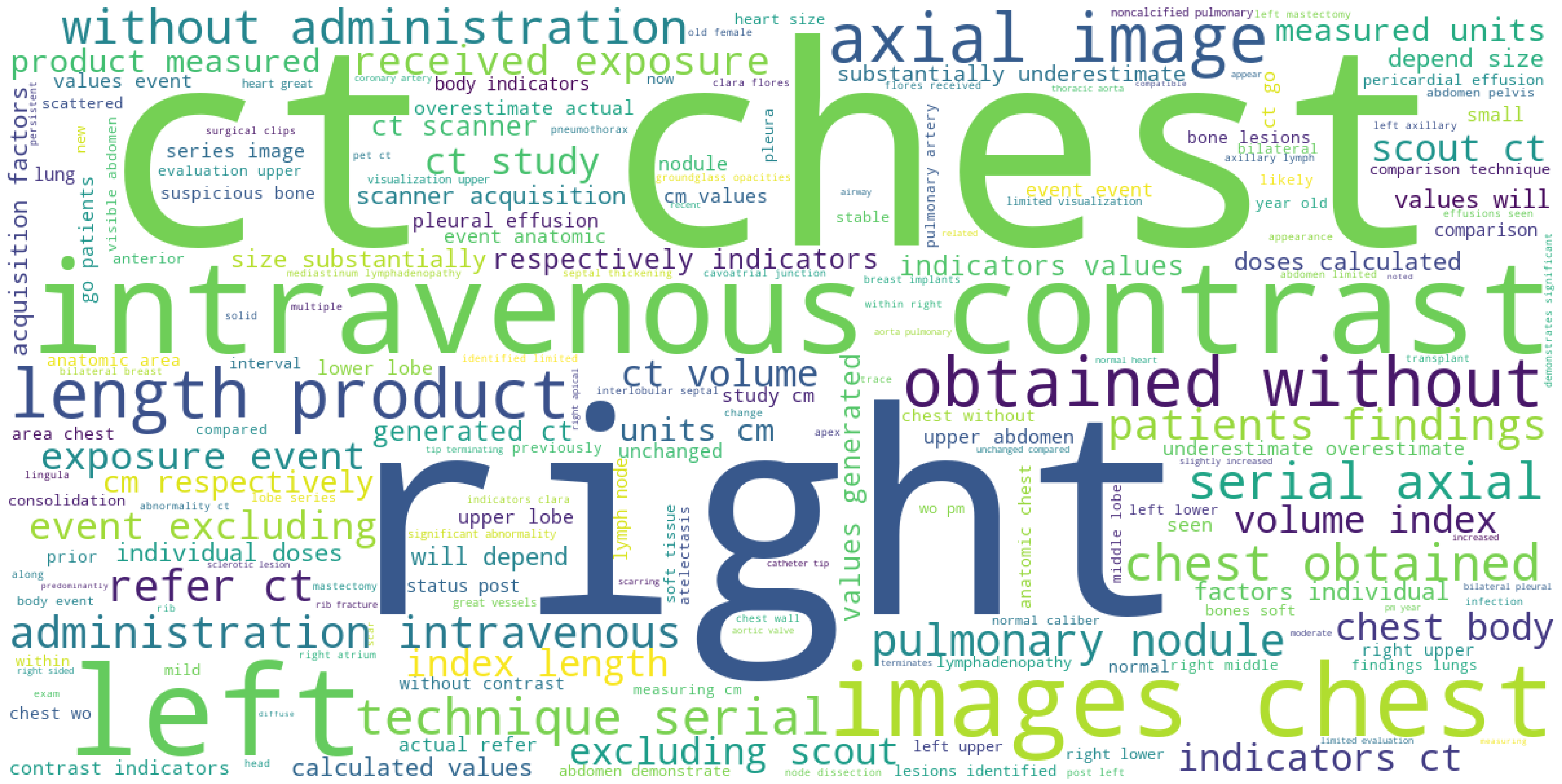
NOTES



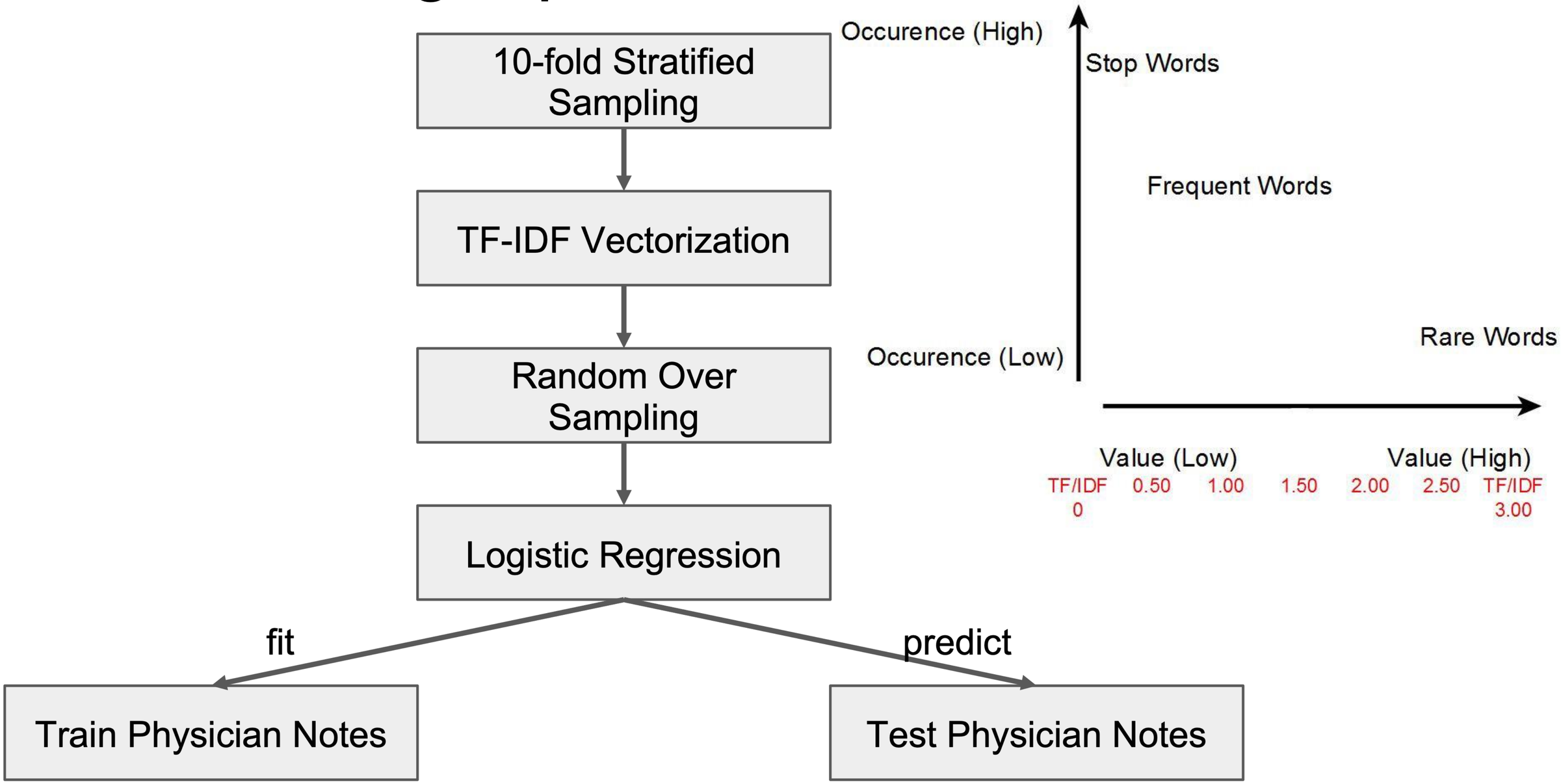
PATHOLOGY



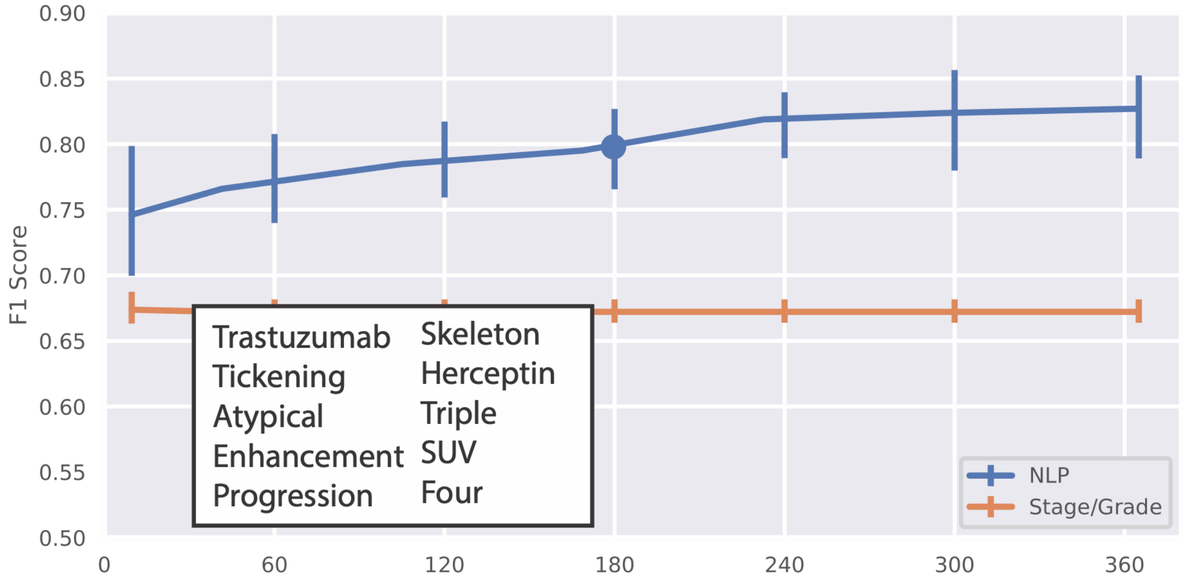
IMAGING

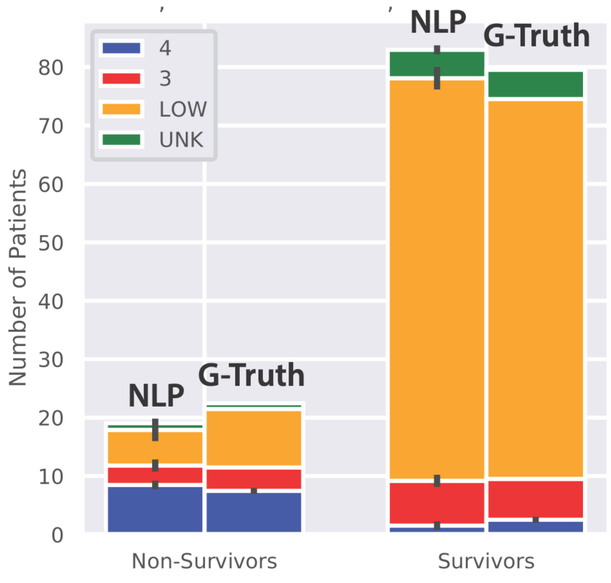


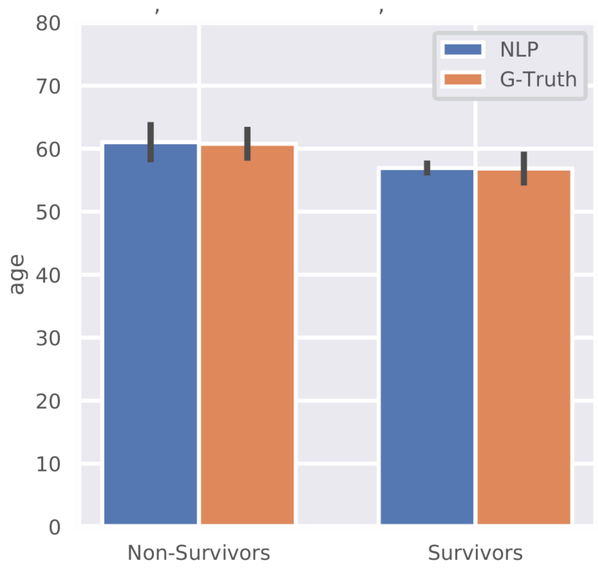
Training Pipeline and Modelling



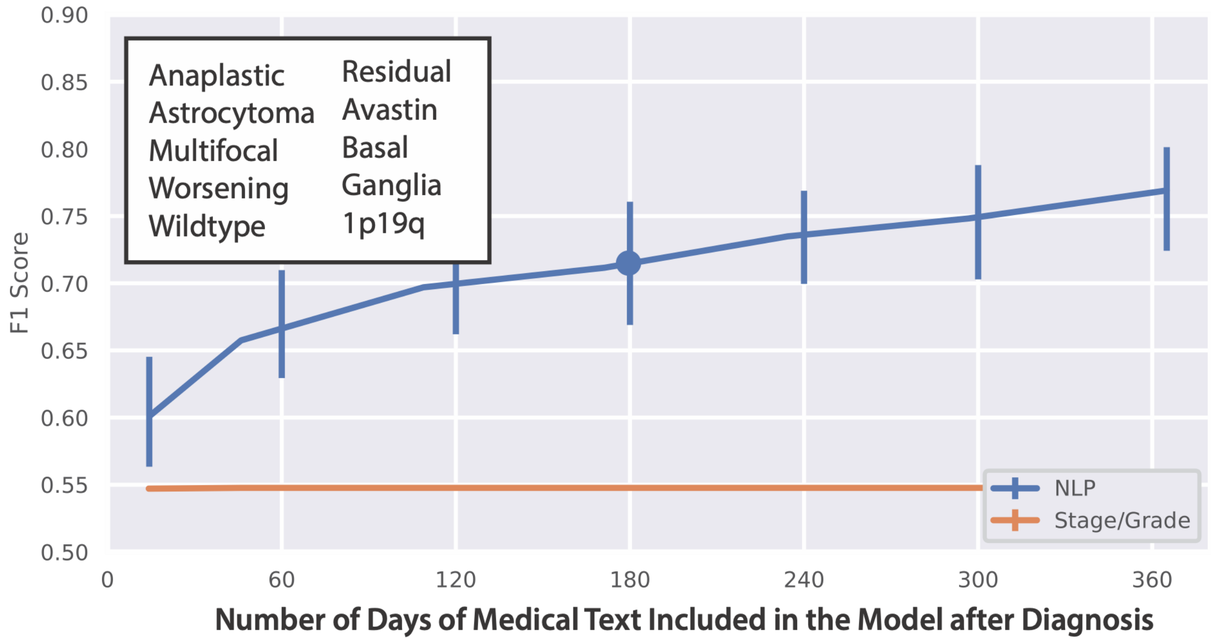
Breast: Predicting 5 yrs Survival, Train 401/Test 102

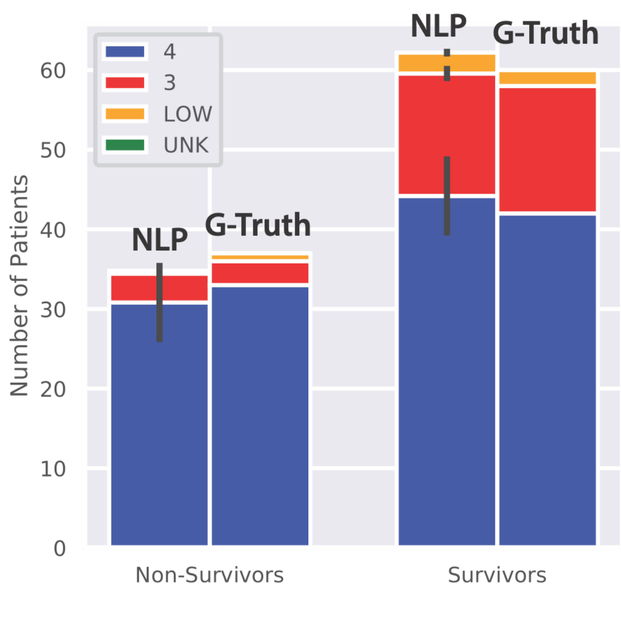


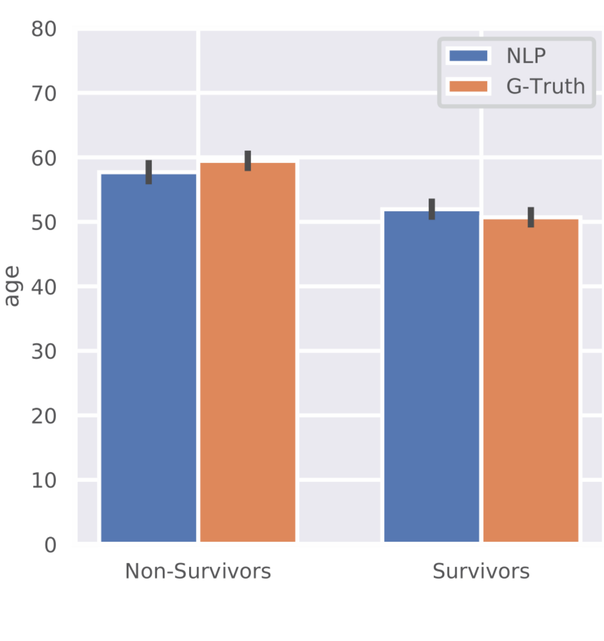




Glioma: Predicting 14 months Survival, Train 378/Test 107







In Summary

- The needs for medical informatics are diverse (administrative, QI, research).
- Data governance and aggregation should be planned carefully.
- We have created a MEDomics framework:
  - **To explore current and new hypotheses from real-world data.**
  - **To develop novel algorithms and clinical tools.**
  - **To take a first step towards self-cognizant and responsive EHR/OIS.**