

The Cancer Imaging Archive (TCIA)

diagnostic image collections case-matched to

The Cancer Genome Atlas (TCGA)

<http://cancerimagingarchive.net/>

John Freymann

Justin Kirby

Carl Jaffe

Fred Prior

Leidos Biomedical, Inc. (NCI Cancer Imaging Program)

Leidos Biomedical, Inc. (NCI Cancer Imaging Program)

Boston University School of Medicine

Washington University St Louis

Summary

1. TCIA provides access to clinical imaging of TCGA cancers
2. Imaging data is a multi-dimensional source of quantifiable information that can complement the TCGA omics
 - a. Spatial and temporal heterogeneity
 - b. Bio-functional properties
3. Ongoing research is providing promising results

TCGA Imaging Data Collection

- Contacted Radiology Depts at TSS supplier institutions
- Arranged PHI compliant transfer of pre-op clinical diagnostic images (MR, CT, NM)
- Images curated by robust HIPAA compliant automatic software with human oversight
- Publicly accessible following TCGA model



**LOG IN TO
THE CANCER
IMAGING ARCHIVE**

The Cancer Imaging Archive (TCIA) is a large and growing archive of medical images of cancer which are available to download for research purposes. [Learn more...](#)

PHOTO COURTESY OF SIEMENS HEALTHCARE

- 1
- 2
- 3
- 4
- 5

ABOUT US

- About The Cancer Imaging Archive (TCIA)
- About the Cancer Imaging Program (CIP)
- About Washington University (WUSTL)
- What's New with TCIA

FOR RESEARCHERS

- Getting Access to the Images
- The Image Collections
- Related Publications
- Research Projects

IMAGE SUBMISSIONS

- Requesting Permission to Upload Your Data
- What to Expect as an Image Provider
- Getting Help with Submissions
- De-identification Knowledge Base

Publicly available cases as of May 2014

BLCA - 6

KIRP - 20

BRCA - 122

LGG - 82

COAD - 5

LIHC - 65

GBM - 256

LUAD - 64

HNSC - 27

LUSC - 3

KICH - 15

OV - 50

KIRC - 212

PRAD, READ, THCA < 5

IMAGE ARCHIVE



GENE ARCHIVE

Search Results (by Sub

Subjects checked can be added to your basket.

Legend:

- A Available
- P Pending
- N Not Available
- Not Applicable

*Protected data

CHECK ALL **UNCHECK ALL**

	Collection ID	Subject ID
Show Studies	TCGA-GBM	TCGA-02-0003
Show Studies	TCGA-GBM	TCGA-02-0006
Show Studies	TCGA-GBM	TCGA-02-0009
Show Studies	TCGA-GBM	TCGA-02-0011
Show Studies	TCGA-GBM	TCGA-02-0027
Show Studies	TCGA-GBM	TCGA-02-0033

Batch/Sample	Level			1	2	3	1
Batch 1	TCGA-02-0001-01	A	A	A	A	A	A
	TCGA-02-0003-01	A	A	A	A	A	A
	TCGA-02-0006-01	A	A				A
	TCGA-02-0007-01	A	A	A	A	A	A
	TCGA-02-0009-01	A	A	A	A	A	A
	TCGA-02-0010-01	A	A	A	A	A	A
	TCGA-02-0011-01	A	A	A	A	A	A
	TCGA-02-0014-01	A	A	A	A	A	A
	TCGA-02-0021-01	A	A	A	A	A	A
	TCGA-02-0024-01	A	A	A	A	A	A

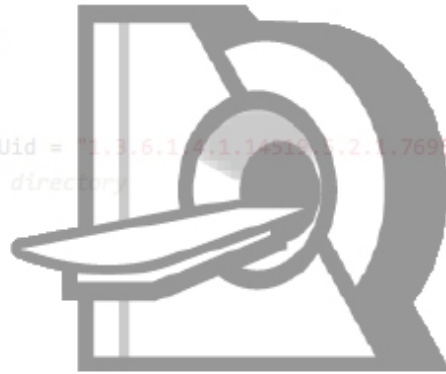
							Exp-Gene
XML	Biotab	HT_HG-U133A					



```

104 except urllib2.HTTPError, err:
105     print "Error executing program:\nError Code: ", str(err.code) , "\nMessage: " , err.re
106
107 Get images in your apps
108 with the new
109 TCIA REST API
110 response = tcia_client.get_image(seriesInstanceId = "1.3.6.1.4.1.14519.2.1.7695.400
111 # tcia.get_image
112 response.getcode() == 200.
113     print "\n" + str(response.info())
114     bytesRead = response.read()
115     fout = open("images.zip", "wb")
116     fout.write(bytesRead)
117     fout.close()

```



LOG IN TO
THE CANCER
IMAGING ARCHIVE

Develop imaging apps that leverage TCIA using the new REST API. Examples in Python and Java are available to help you get started.

[Learn more...](#)

- 1
- 2
- 3
- 4
- 5

 **ABOUT US**

- About The Cancer Imaging Archive (TCIA)
- About the Cancer Imaging Program (CIP)
- About Washington University (WUSTL)
- What's New with TCIA

 **FOR RESEARCHERS**

- Getting Access to the Images
- The Image Collections
- Related Publications
- Research Projects

 **IMAGE SUBMISSIONS**

- Requesting Permission to Upload Your Data
- What to Expect as an Image Provider
- Getting Help with Submissions
- De-identification Knowledge Base

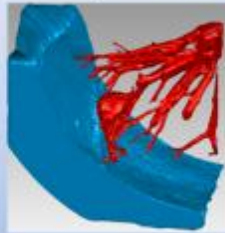


Select your Images

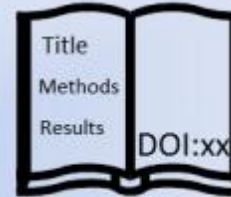
Request DOI

 Citation linking for datasets

Analyze Data



Publish with DOI



Support Open Science:
Publish your data with a
TCIA DOI

Retrieve Data using DOI



LOG IN TO
THE CANCER
IMAGING ARCHIVE

TCIA has the ability to create [Digital Object Identifier \(DOI\)](#) linked to subsets of TCIA data, which authors may use as data citations in their scholarly papers. [Learn more...](#)

- 1
- 2
- 3
- 4
- 5



ABOUT US

- About The Cancer Imaging Archive (TCIA)
- About the Cancer Imaging Program (CIP)
- About Washington University (WUSTL)
- What's New with TCIA



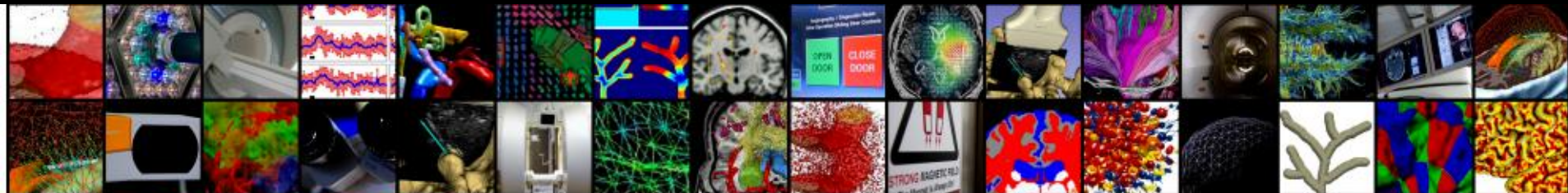
FOR RESEARCHERS

- Getting Access to the Images
- The Image Collections
- Related Publications
- Research Projects



IMAGE SUBMISSIONS

- Requesting Permission to Upload Your Data
- What to Expect as an Image Provider
- Getting Help with Submissions
- De-identification Knowledge Base



MICCAI 2014 BOSTON

- Home
- Workshops, Challenges, Tutorials
- Paper Submission
- Paper Reviewing
- Important Dates
- Program
- Organization
- Sponsorship
- Registration
- Local Information
- Graphics and Credits
- Contact

Updates

April 24, 2014: If you were invited to submit a rebuttal, when you login into the system, it will show the status of your paper as "not accepted." Please don't panic and take it as "not accepted yet." We had to make the status visible to enable the authors of the early accepted papers to upload the final version of their papers. Also, your rebuttals are in the system. You can't see anything because we switched the system into the discussion mode. Everything will be available again once the decisions are made.

April 15, 2014: Please note that the final versions of all early accepted papers are due on **May 15, 2014, 23:59 PST**. We will post instructions for submitting the camera ready copy this week.

April 14, 2014: Early accept/reject decisions have been emailed to contact authors. The reviews have been posted for all papers. The rebuttals can be submitted until Monday **April 21, 2014, 23:59 PST**. Early accepted/rejected papers cannot submit rebuttals; only submissions that were invited to submit a rebuttal will have a link to rebuttal on their page.

April 9, 2014: The review deadline was April 8. We are following up on late reviews and will post as soon as we have all reviews in. We will allow sufficient time for rebuttals.

March 11, 2014: This year, we are introducing a new [MICCAI Educational Challenge](#)!

March 10, 2014: The [Workshop, Challenge and Tutorial Program](#) has been posted. We will update the program with links to the workshops' web pages as they become available.

February 28, 2014: The paper submission deadline is firm; there will be no extensions. We will keep the system open for a couple of hours past the deadline, to make sure all the last-minute uploads are successful.

February 27, 2014: To be considered for the MICCAI Young Scientist Award and for the student travel awards, the first

September 14-18, 2014

Quick Links:

[Educational Challenge](#)

[Workshops/Challenges/Tutorials](#)

[Become a sponsor](#)

[MICCAI facebook page](#)

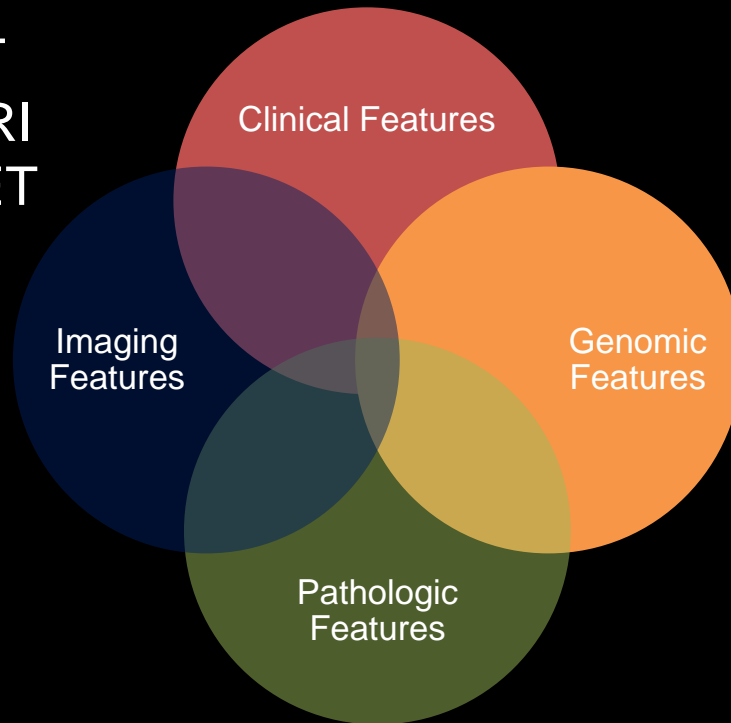
SIEMENS

The Cancer Imaging Archive: Imaging-Genomics through Open Science



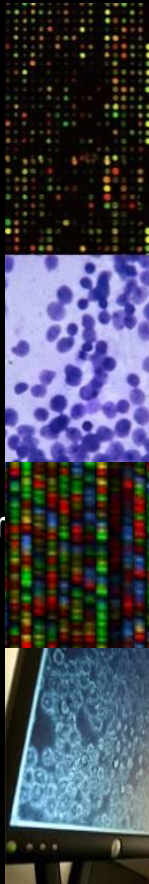
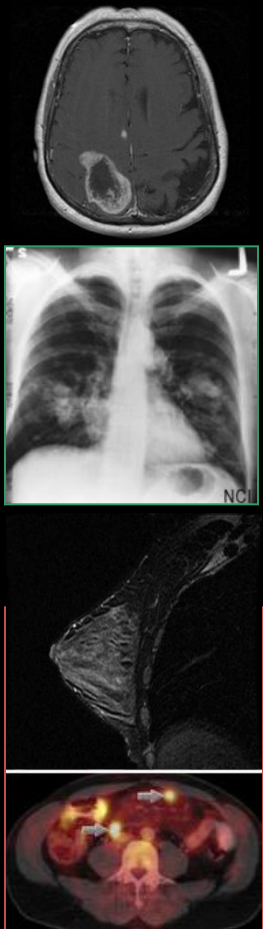
Data types

- CT
- MRI
- PET



Data types

- Clinical diagnosis
- Treatment history
- Histologic diagnosis
- Pathologic report/images
- Tissue anatomic site
- Surgical history
- Gene expression
- RNA sequence
- Chromosomal copy number
- Loss of heterozygosity
- Methylation patterns
- miRNA expression
- DNA sequence
- RPPA (protein)
- Subset for Mass Spec



TCGA clinical images at <http://cancerimagingarchive.net>

clinical imaging contribution to data set:

- Non-invasive
- Provides surrogate endpoints (e.g. when there's long survival curves)
- Temporal data: can re-image over time, - tumor tissue acquisition is usually single time
- Gives information about tumor heterogeneity (biopsy location)
- Tissue is a finite resource - imaging supports unlimited re-use

clinical imaging analysis:

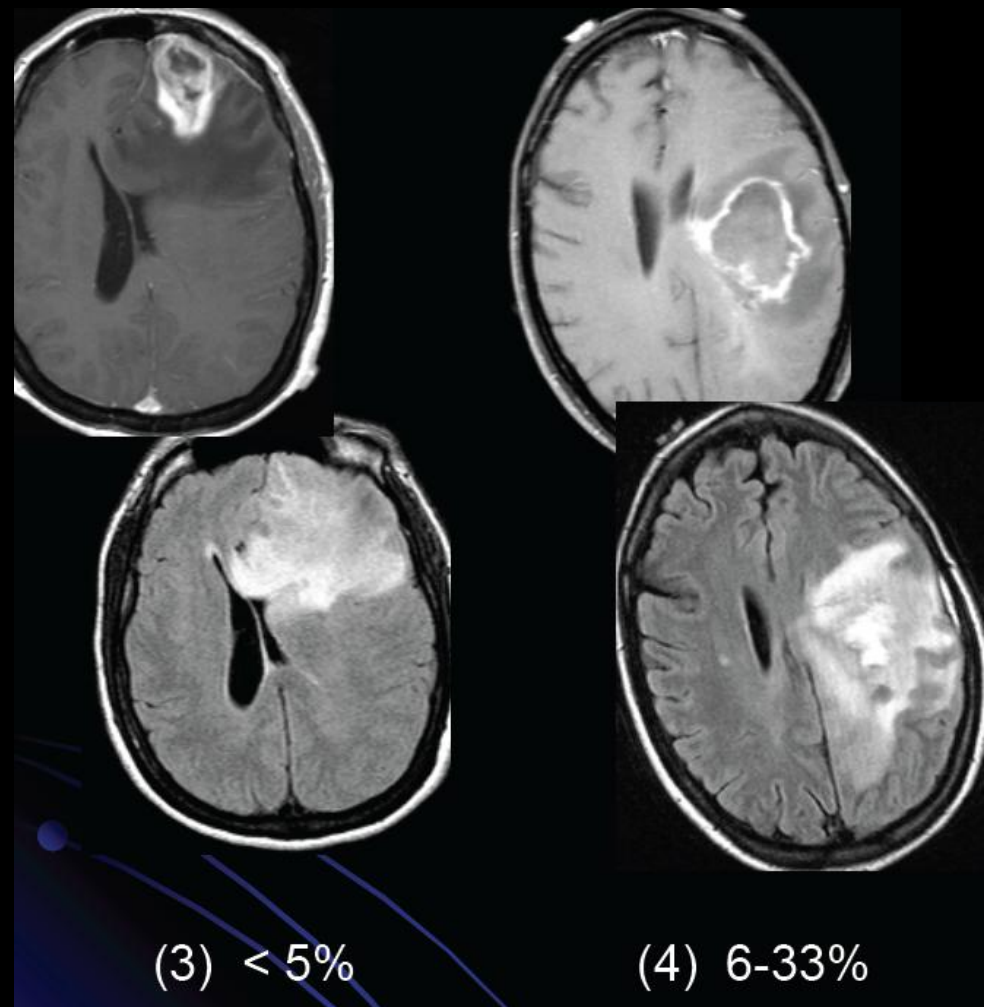
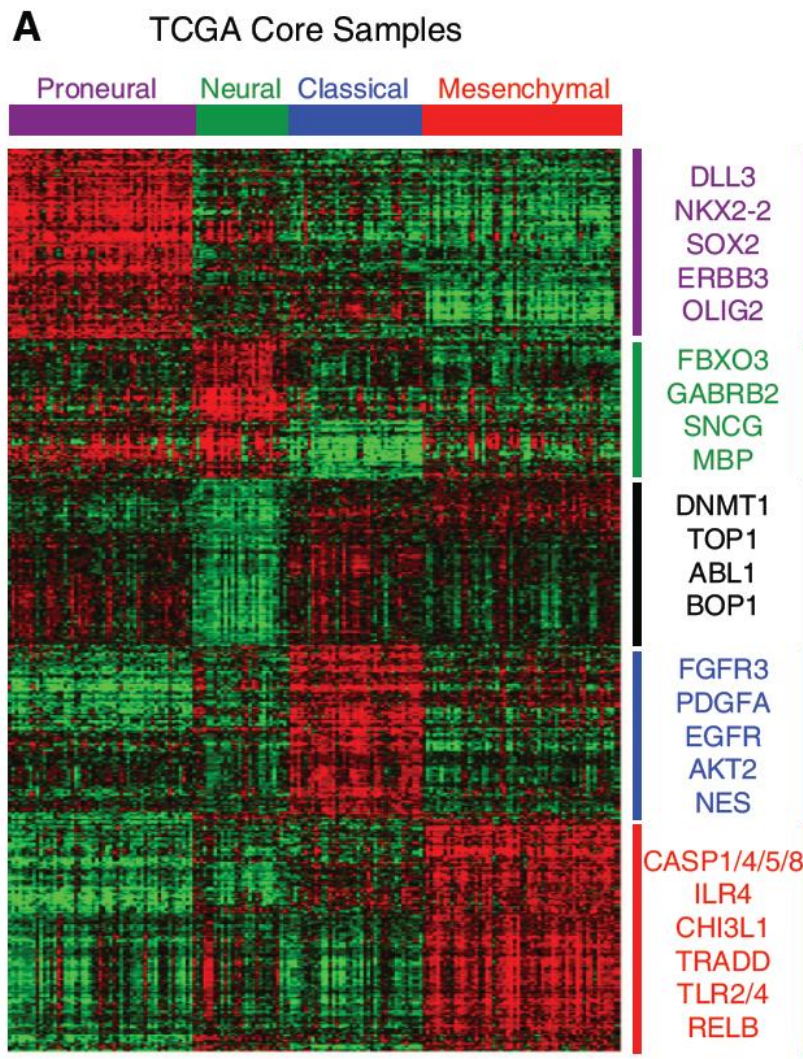
- Tumor anatomic morphology
 - shape, texture
 - Scored with reproducibility by multiple-experts
 - Quantitative by computer image processing
- Contrast dynamics
 - perfusion blood flow; diffusion
- Tumor matrix and its temporal response
- Molecular insights (functional imaging)
 - Nuclear agent tagged molecules, e.g. PET

Exploring tumor complexity clinical implications

- genotypic – epigenetic – proteomic diversity
- clinical imaging pleomorphisms (MR, NM, CT)
- variations in therapeutic response / survival

- Can imaging contribute to prognostic or predictive gene-driven disease “signatures” ?

GENOTYPE \leftrightarrow PHENOTYPE exploring relationships of tumor neo-vascularity ?



MR Imaging Predictors of Molecular Profile and Survival: Multi-institutional Study of the TCGA Glioblastoma Data Set¹

David A. Gutman, MD, PhD
Lee A. D. Cooper, PhD
Scott N. Hwang, MD, PhD
Chad A. Holder, MD
JingJing Gao, PhD
Tarun D. Aurora, BS
William D. Dunn, Jr, BS
Lisa Scarpace, MS
Tom Mikkelsen, MD
Rajan Jain, MD
Max Wintermark, MD, MAS
Manal Jilwan, MD
Prashant Raghavan, MD
Erich Huang, PhD
Robert J. Clifford, PhD
Pattanasak Mongkolwat, PhD
Vladimir Kleper, BS
John Freymann, BA
Justin Kirby, BS
Pascal O. Zinn, MD

Purpose:

To conduct a comprehensive analysis of radiologist-made assessments of glioblastoma (GBM) tumor size and composition by using a community-developed controlled terminology of magnetic resonance (MR) imaging visual features as they relate to genetic alterations, gene expression class, and patient survival.

Materials and Methods:

Because all study patients had been previously deidentified by the Cancer Genome Atlas (TCGA), a publicly available data set that contains no linkage to patient identifiers and that is HIPAA compliant, no institutional review board approval was required. Presurgical MR images of 75 patients with GBM with genetic data in the TCGA portal were rated by three neuroradiologists for size, location, and tumor morphology by using a standardized feature set. Interrater agreements were analyzed by using the Krippendorff α statistic and intraclass correlation coefficient. Associations between survival, tumor size, and morphology were determined by using multivariate Cox regression models; associations between imaging features and genomics were studied by using the Fisher exact test.

Radiogenomics of Clear Cell Renal Cell Carcinoma: Associations between CT Imaging Features and Mutations

Christoph A. Karlo, MD, Pier Luigi Di Paolo, MD, Joshua Chaim, DO, A Ari Hakimi, MD, Irina Ostrovnya, PhD, Paul Russo, MD, FACS, Hedvig Hricak, MD, PhD, Dr(hc), Robert Motzer, MD, James J. Hsieh, MD, PhD, Oguz Akin, MD

From the Genitourinary Imaging Group, Department of Radiology (C.A.K., P.L.D.P., J.C., H.H., O.A.), Urology Service, Department of Surgery (A.A.H., P.R.), Human Oncology & Pathogenesis Program (A.A.H., J.J.H.), Department of Epidemiology and Biostatistics (I.O.), and Department of Medicine, Genitourinary Oncology Service (R.M., J.J.H.), Memorial Sloan-Kettering Cancer Center, 1275 York Ave, Radiology Academic Offices, Room C278, New York, NY 10065.

DOI: <http://dx.doi.org/10.1148/radiol.13130663>

Abstract

Full Text

Figures

References

Supplemental Materials

Cited by

PDF

Purpose

To investigate associations between computed tomographic (CT) features of clear cell renal cell carcinoma (RCC) and mutations in *VHL*, *PBRM1*, *SETD2*, *KDM5C*, or *BAP1* genes.

Materials and Methods

The institutional review board approved this retrospective, hypothesis-generating study of 233 patients with clear cell RCC and waived the informed consent requirement. The study was HIPAA compliant. Three radiologists independently reviewed pretreatment CT images of all clear cell RCCs without knowledge of their genomic profile. One radiologist measured largest diameter and enhancement parameters of each clear cell RCC. Associations between CT features and mutations in *VHL*, *PBRM1*, *SETD2*, *KDM5C*, and *BAP1* genes were tested by using the Fisher exact test. Associations between mutations and size and enhancement were assessed by using the independent *t* test. Interreader agreement was calculated by using the Fleiss κ .

Define image features, possible answers

VASARI_4_0_featurekey-2013-10-01 - Microsoft Word

Home Insert Page Layout References Mailings Review View

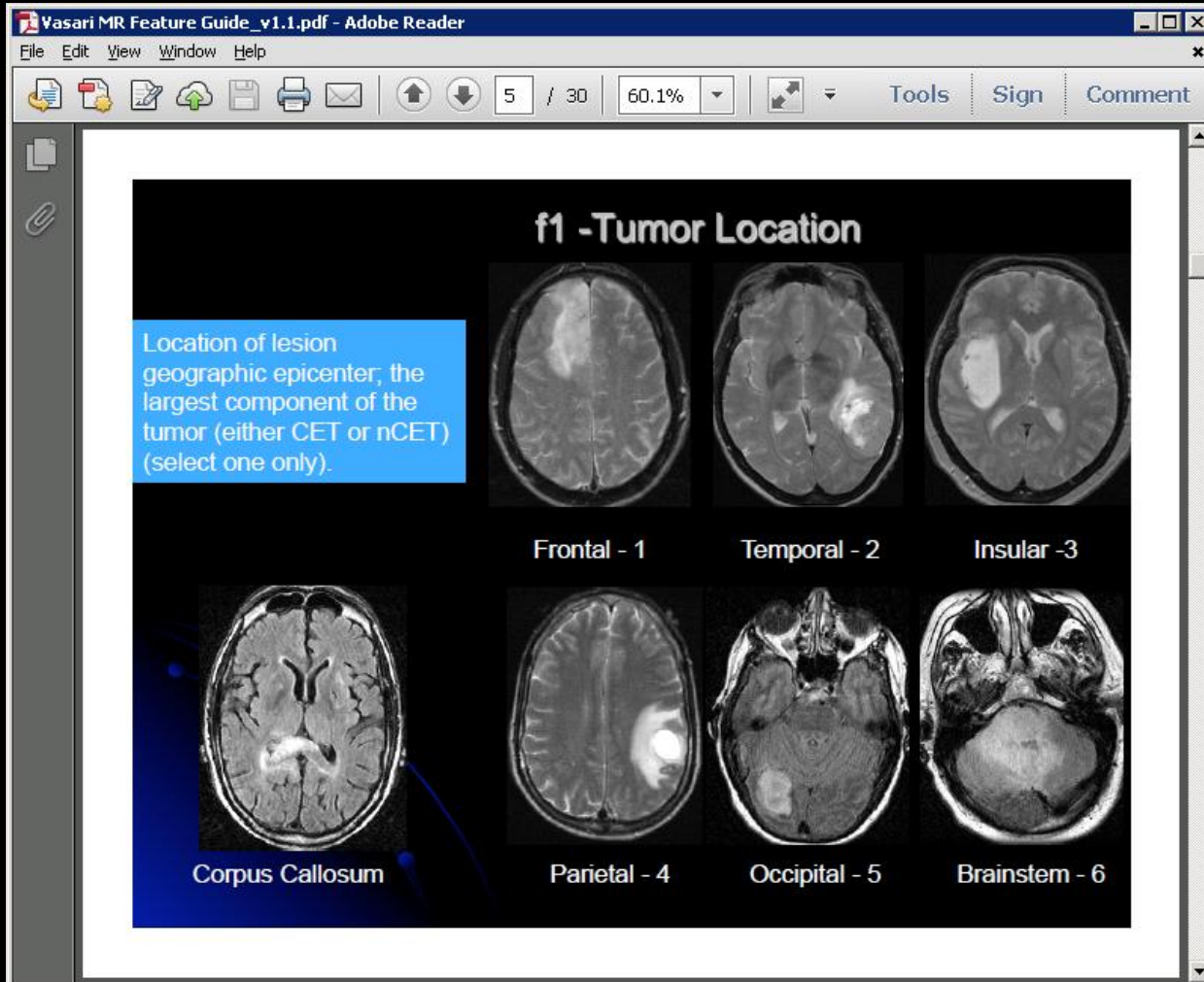
Clipboard Font Paragraph Styles Editing

VASARI 4.0 TCGA *inSilico* MR Feature KEY 09/24/2013
 Baseline Glioma MR tumor assessment to include potential features for low grade glioma (LGG).

Feature number/name	Description	Options
F0 Image QA	Quality assurance pre-check. Inventory of required image series including FLAIR/T2 and pre/post contrast T1 WI.	<ul style="list-style-type: none"> Post biopsy (requires adjudication) Post-op study (disqualified) No contrast injected (disqualified) No T2/FLAIR images (disqualified)
F1 Tumor Center	Location of lesion geographic center; the largest component of the tumor (either CET or nCET). (multiple selections acceptable – choose up to two)	<ul style="list-style-type: none"> Frontal lobe Temporal lobe Parietal lobe Occipital lobe Insula Brainstem Cerebellum Lentiform nucleus Caudate nucleus Thalamus Corpus callosum
F34 Tumor Origin	Propose where you think the tumor originated from using the following grades. Periventricular (most likely arising from periventricular/subventricular zone); Cyral (most likely arising from the cortex or adjacent cyral subcortical white matter zone); Deep white matter (most likely arising deep to the U fiber region but not in the periventricular zone). Use not applicable if none of these feature criteria apply. (multiple selections acceptable)	<ul style="list-style-type: none"> Periventricular Cyral Deep white matter Other

Page: 1 of 8 Words: 1,986 70%

Create training guides



Generate structured data capture templates

The screenshot displays the AIM Template Builder 2.0 interface. The main window is titled "AIM Template Builder" and shows the configuration for a template named "VASARI 4.0". The interface is divided into several sections:

- Template Groups:** A list of templates including LIDC CT, MarkupOnly-6, RSNA MR-Cardiac, RSNA Template - MR, TCGA-BRCA CRF, TCGA-OV - Round 1, VASARI_9, VASARI_caDSR, and VASARI 4.0.
- Template Components:** A hierarchical tree structure showing the components of the VASARI 4.0 template. The root is "VASARI 4.0 (ver 4.0 Baseline Glioma MR tumor assessment to include...)", which branches into "Tumor (min:1 max:1)". Under "Tumor", there are three main categories, each with multiple answer choices:
 - Imaging Observation Characteristic: Image QA**
 - Answer Choice: Post biopsy (requires adjudication)
 - Answer Choice: Post-op study (disqualified)
 - Answer Choice: No Contrast Injected (disqualified)
 - Answer Choice: No T2/FLAIR images (disqualified)
 - Imaging Observation Characteristic: Tumor Center**
 - Answer Choice: frontal lobe
 - Answer Choice: temporal lobe
 - Answer Choice: parietal lobe
 - Answer Choice: occipital lobe
 - Answer Choice: insula
 - Answer Choice: brainstem
 - Answer Choice: cerebellum
 - Answer Choice: lentiform nucleus
 - Answer Choice: caudate nucleus
 - Answer Choice: thalamus
 - Answer Choice: corpus callosum
 - Imaging Observation Characteristic: Tumor Origin**
 - Answer Choice: Periventricular
 - Answer Choice: Gyral
 - Answer Choice: Deep white matter
 - Answer Choice: Other
 - Imaging Observation Characteristic: Side of Tumor Center**
 - Answer Choice: Right
 - Answer Choice: Center/Bilateral
- Properties Panel:** A form on the right side of the window containing the following fields:
 - Name*: VASARI 4.0
 - Description*: Baseline Glioma MR tumor assessment to include potential features for low grade glioma (LGG).
 - Authors: Justin Kirby
 - Modality: MR
 - Preceding Annotation:
 - UID*: 2.25.2775549064325515900743877
 - Date: 2013-08-26
 - Version*: 4.0
 - Lexicon Term:
 - Code Meaning*: VARSARI
 - Code Value*: 99PRI_TC_0A23
 - Schema Designator*: 99PRI_TemplateCode
 - Schema Version: 1.0
 - Select Term button
- Save Button:** A green "Save" button at the bottom right.



Radiologists use the template to assess the images

TCGA-CS-5395 - TCGA-CS-5395 [AIM on ClearCanvas Workstation 3.5.0.1]

File Edit View Tools Annotation Help

TCGA-CS-5395 TCGA-CS-5395

TCGA-CS-5395
M
043Y

BRAIN_CONTRAST/FLAIR

04-Oct-1998
Acq: 14:52:13.351
Se: 3/1
m: 28/20
Loc: H39.8

3.49 cm
2.29 cm

512 x 140
ET: 1
TR: 9000.00 ms
TE: 105.00 ms
CP Head
Thk: 5.0 mm
Zoom: 2.78x
W:1128 L:535 (Auto)

1.5T
mrscan
MAGNETOM VISION
DFOV: 22.0 x 22.0 cm

AIM Annotation

AIM Template: VARSARI(99PRI_TC_0A23)

Image QA

- Post biopsy (requires adjudication)
- Post-op study (disqualified)
- No Contrast Injected (disqualified)
- No T2/FLAIR images (disqualified)

Tumor Center
Select at least 1 and less than 3

- frontal lobe
- temporal lobe
- parietal lobe
- occipital lobe
- insula
- brainstem
- cerebellum
- lentiform nucleus
- caudate nucleus
- thalamus
- corpus callosum

Tumor Origin
Select at least 1

- Periventricular
- Gyral
- Deep white matter
- Other

Side of Tumor Center

Left

Eloquent Brain
Select at least 1 and less than 5

- No Eloquent Brain
- Speech Motor
- Speech Receptive
- Motor
- Vision

Reset Create Annotation User Info

Export radiologist assessments to flat files

Query API Browser

Project: **TCGA-KIRC**

Description: TCGA KIRC (renal) Research Group

Data Provider: **Round-2**

Description : Renal group round 2 evaluations

GET allxml

GET results

Description :

URL : <https://aime.cancerimagingarchive.net/services/TCGA-KIRC/Round-2/query/results>

Name	Description	Value	Required	Default Value
reviewer		<input type="text"/>	false	
patientID		<input type="text"/>	false	

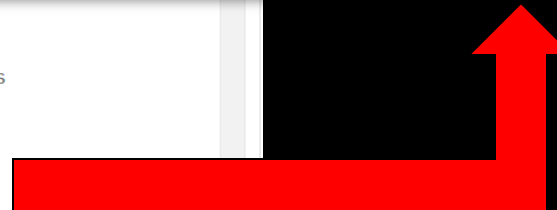
Submit

GET validation

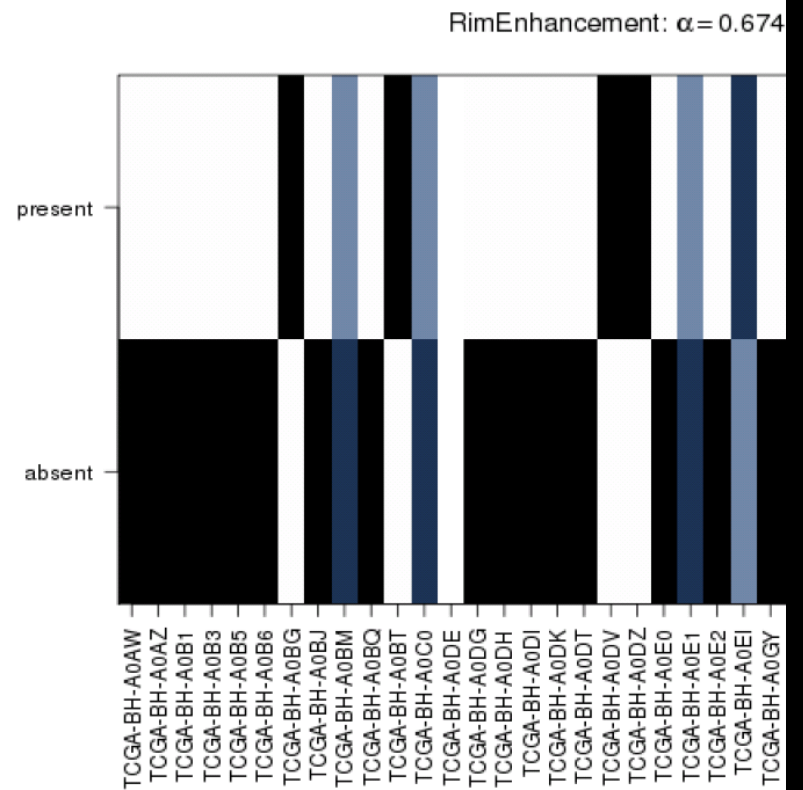
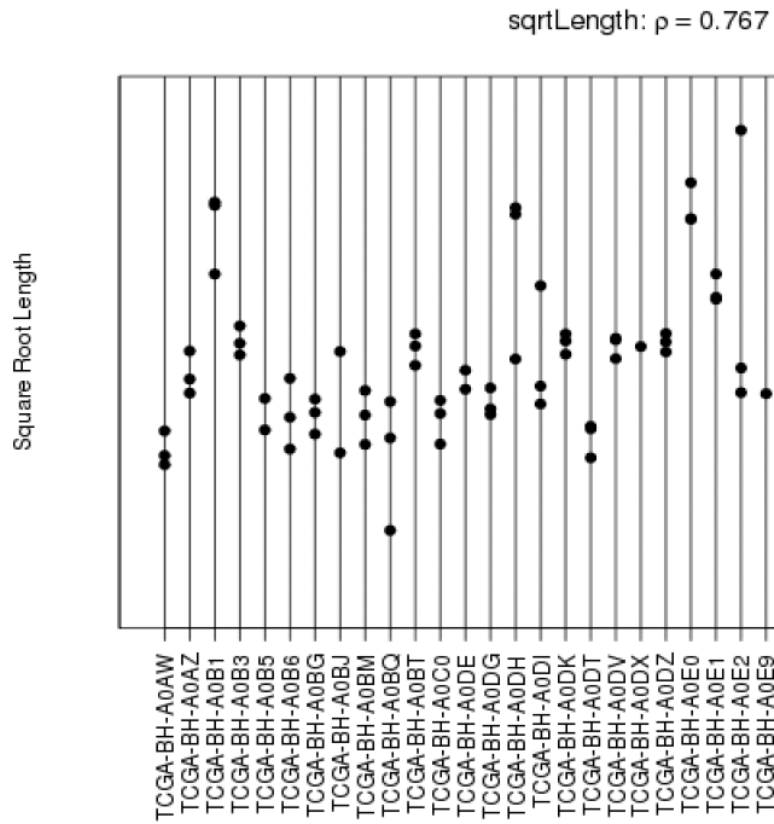
POST xml

DELETE delete

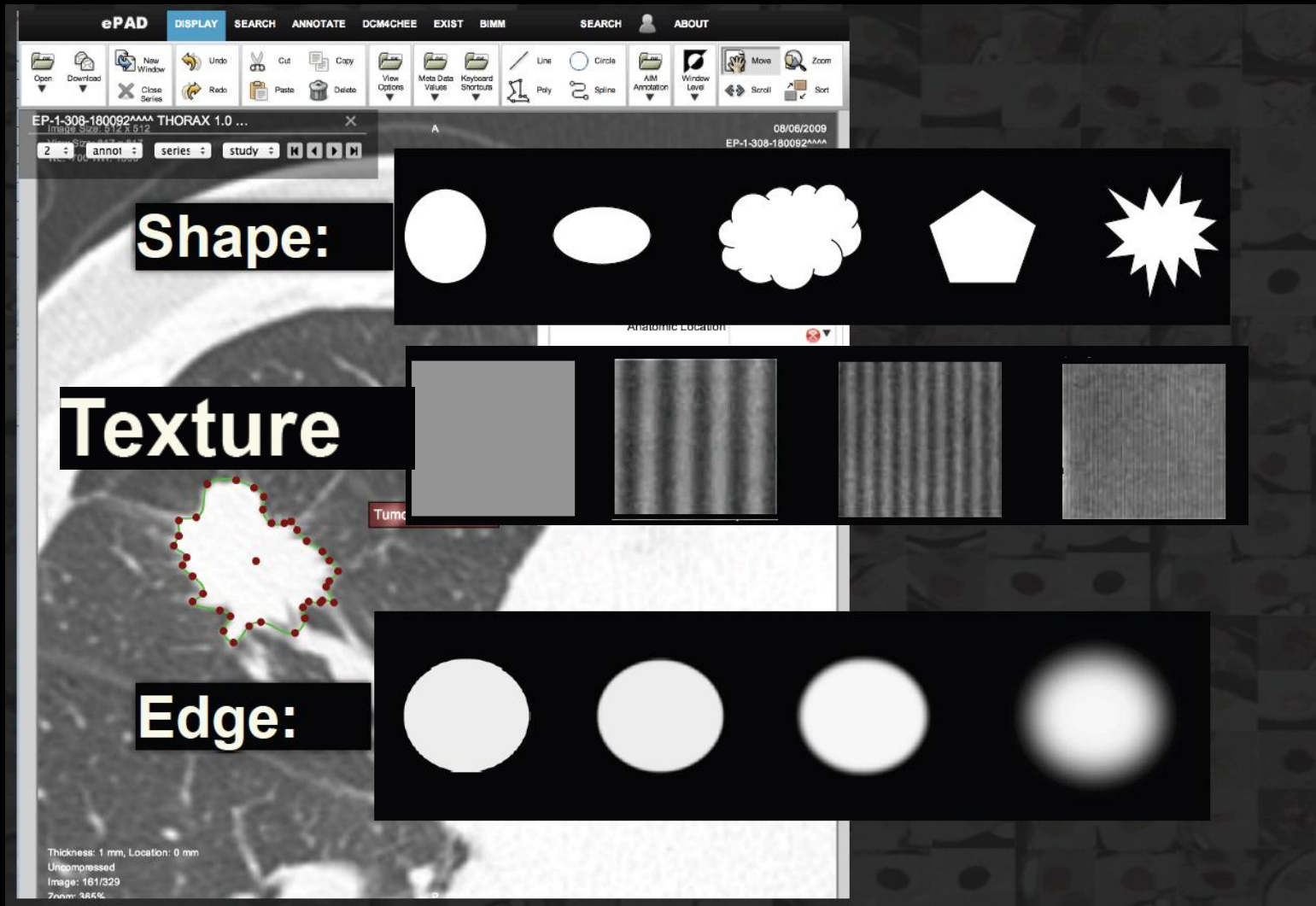
PATIENT_ID	MULTIPLE_TUMORS	TUMORS_LEFT_SIDE	TUMORS_RIGHT_SIDE	MODALITY	CONTRAST_PHASES
TCGA-BP-5010	no	Not Applicable	Not Applicable	CT	Precontrast Nephrographic
TCGA-BP-5010	no	Not Applicable	Not Applicable	CT	Precontrast Nephrographic
TCGA-BP-5168	no	Not Applicable	Not Applicable	CT	Corticomedullary
TCGA-BP-5169	no	Not Applicable	Not Applicable	CT	Corticomedullary
TCGA-BP-5173	no	Not Applicable	Not Applicable	MR	Precontrast Corticomedullary Nephrographic



Reader agreement analysis



Automated feature extraction



Automated feature extraction

Dynamic MRI Explorer === The University of Chicago ===

The interface displays a breast MRI slice (Slice 31) with a red bounding box around a lesion. Below the main image is a graph showing signal intensity over time. The graph has two y-axes: the left axis ranges from 0 to 1200, and the right axis ranges from 0 to 4. The x-axis represents time from 0 to 5. A white line with 'x' markers shows a signal that rises from approximately 200 at time 0 to 800 at time 1, then remains constant. A pink line with '+' markers shows a signal that rises from 0 at time 0 to 4 at time 1, then remains constant. The coordinates (145.25, 115.25) = 529.0000 are displayed above the graph.

Case TCGA-A0-A03V-1

Subtraction(#2 post - pre) #3 Postcontrast

Load... Save... Case 2 Report

Window/Level Bright Contrast

Outline Draw clear

Lesion Seg
Seg. with a box
Start 29 Slc Selection
End 33 Slc Selection
Draw Rect. Save Box
Segment

Seg. with a click
indicate a center Save Center
Segment

Non fat sat. Threshold: 0.2
NC: 2 3 4
select point Show Seg. ON/OFF
clean up seg. undo seg. clean

Computer Analysis Analyze

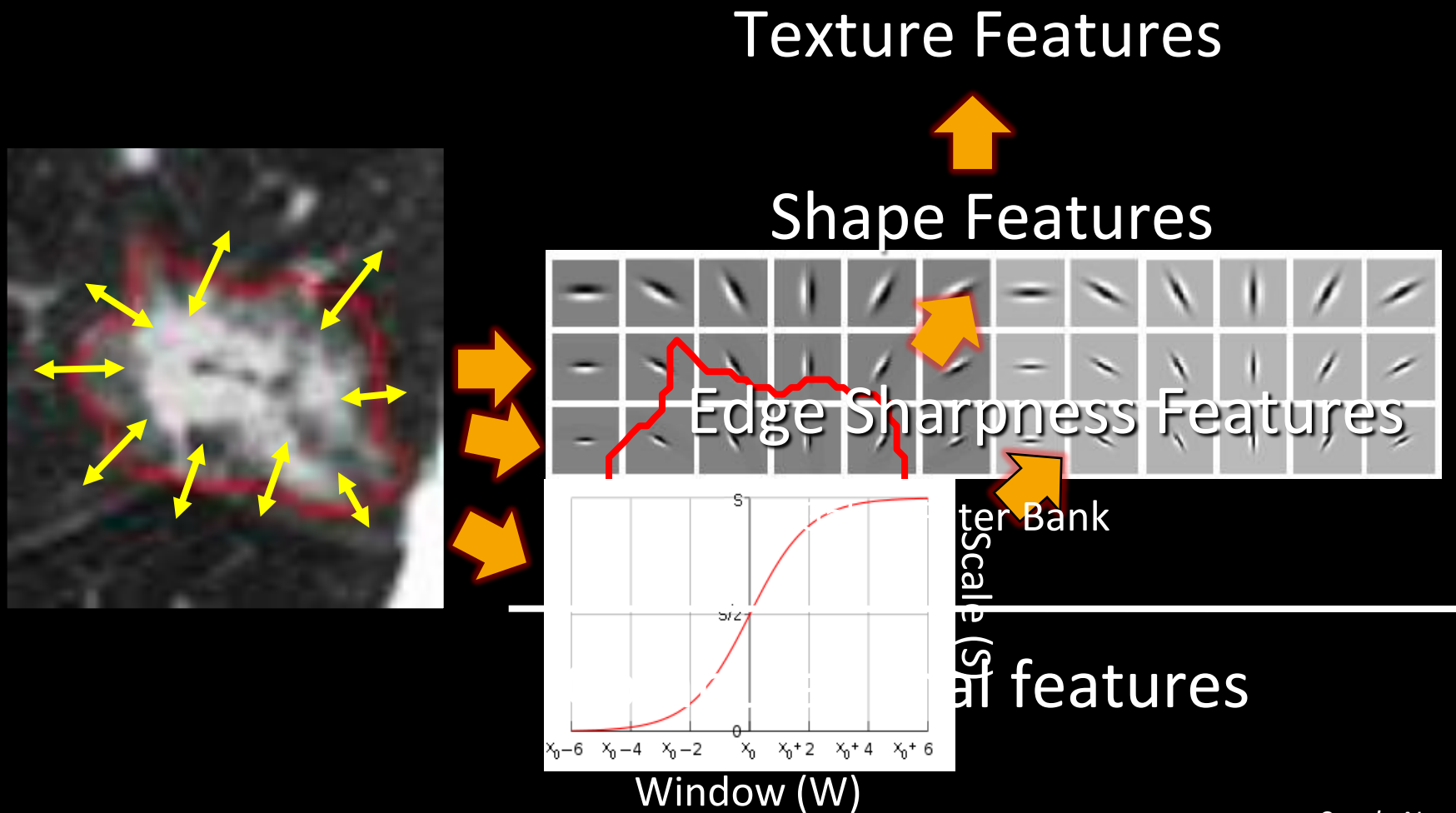
Features
 Prob. of malignancy
 Time to peak
 Variance
 Curve Shape Index
 Washout rate
Show Filter

EXIT

Automated feature extraction

Total features: 36		
	Feature	Feature category
1	Max. enh.	Kinetics feature
2	Time to peak	
3	Uptake rate	
4	Washout rate	
5	Curve shape index	
6	Enhancement at first post-contrast time-point (E1)	
7	Signal enhancement ratio (SER)	
8	Max. var. of enh.	Enhancement-variance kinetics feature
9	Time to peak at maximum variance	
10	Var. incr. Rate	
11	Var. dec. rate	
12	Contrast	Texture feature
13	Correlation	
14	Difference Entropy	
15	Difference Variance	
16	Energy	
17	Entropy	
18	Homogeneity	
19	IMC1	
20	IMC2	
21	Max Corr. Coeff.	
22	Sum Average	
23	Sum Entropy	
24	Sum Variance	
25	Variance	
26	Sphericity	Morphological feature
27	Irregularity	
28	Margin sharpness	
29	Var. of margin sharpness	
30	Variance of RGH	
31	Lesion Volume (mm ³)	Size feature
32	Effective diameter (mm)	
33	Maximum diameter (mm)	
34	Surface area (mm ²)	
35	Surface to volume ratio	
36	Most enhancing volume (mm ³)	

High Throughput Feature Extraction: "Radiomics"



Imaging + genomics – “radiogenomics”

Tissue level

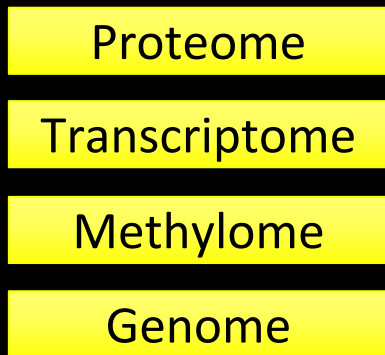


Feature
extraction

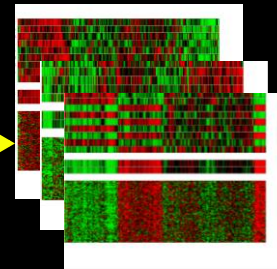
Feature vector
Size $\sim 10^2$

Radiogenomics

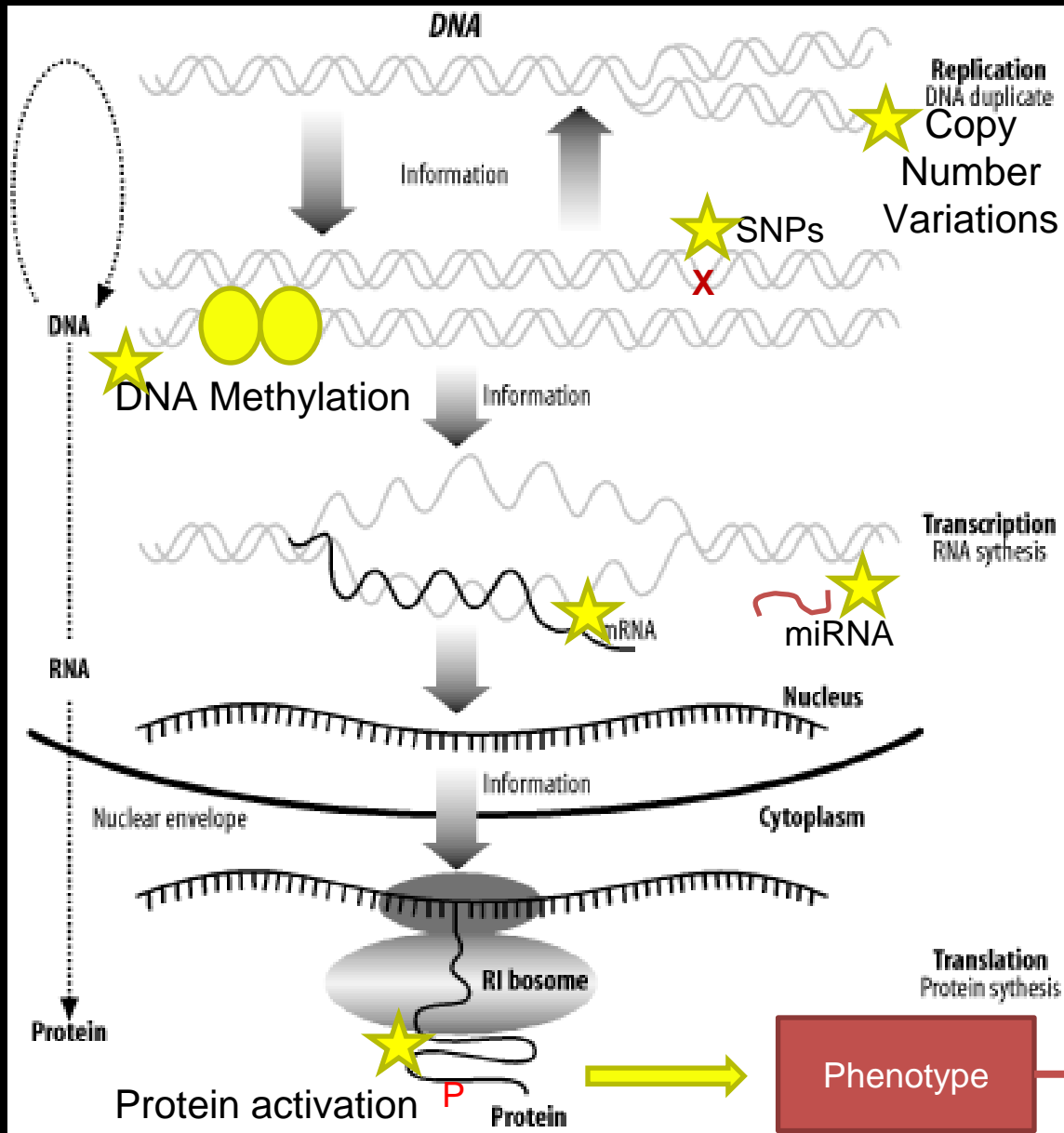
Molecular level



Omics data
integration



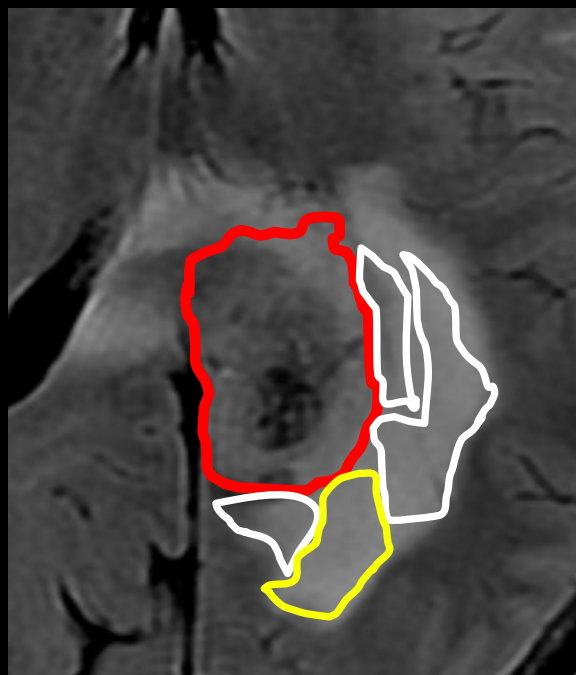
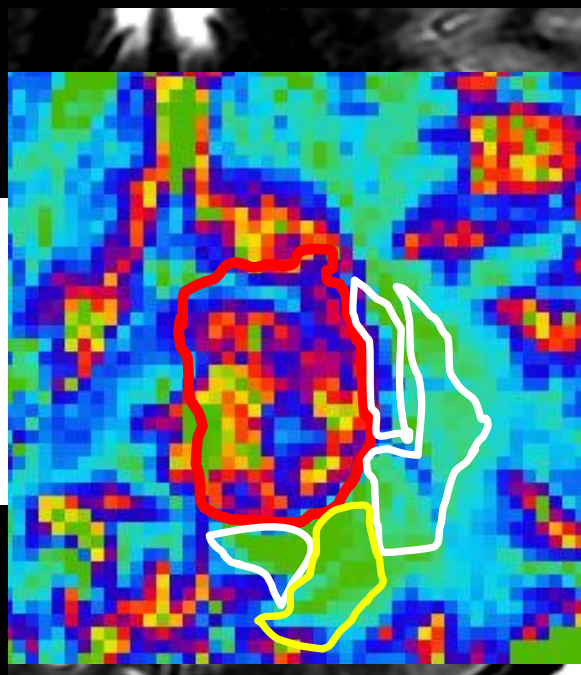
Imaging + genomics – “radiogenomics”



Consider imaging feature as the outcome or phenotype and look for associations with ‘upstream’ molecular elements

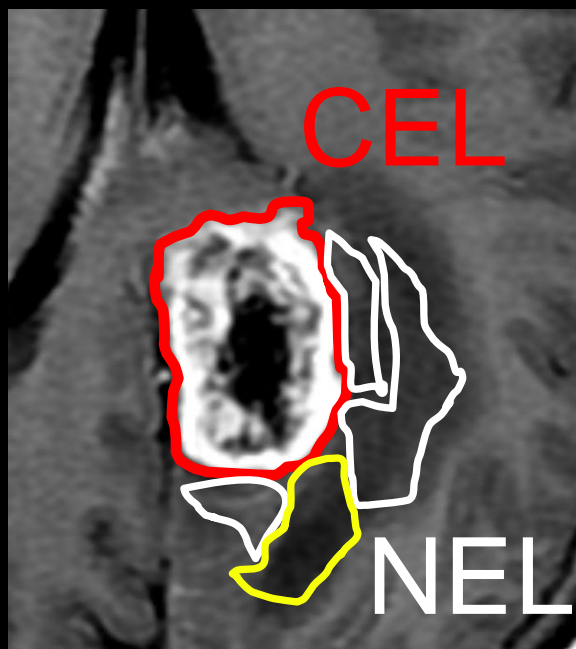
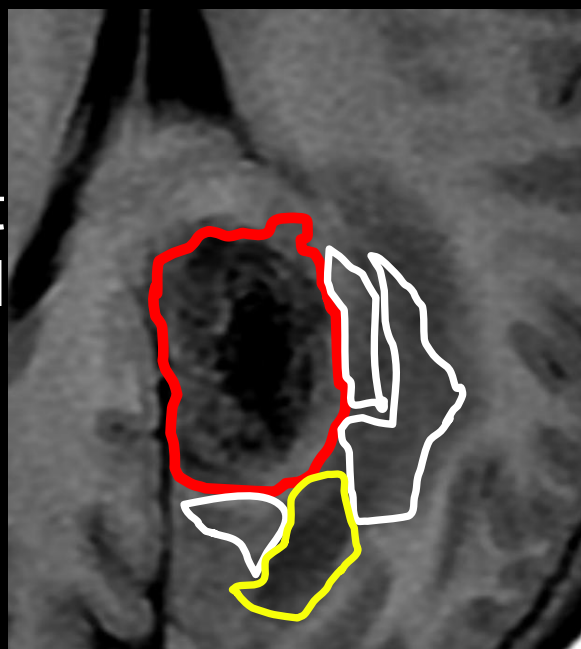
Radiogenomics: Tumor heterogeneity / blood flow

Dynamic
Susceptibility
Contrast-enhanced
T2 Weighted MRI
(DSC T2* MR)
rCBV (green)



FLAIR
MR Image

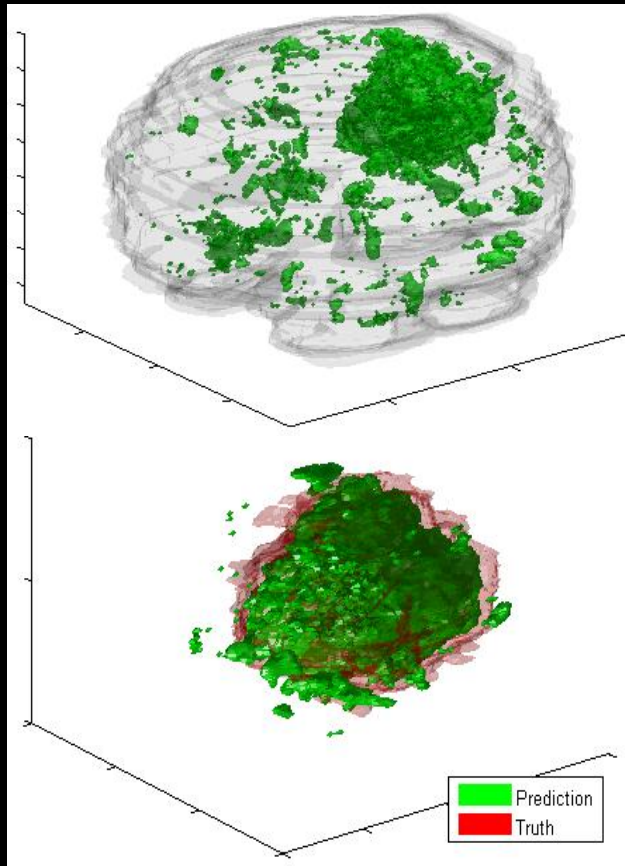
Pre-Contrast
T1 Weighted
MR Image



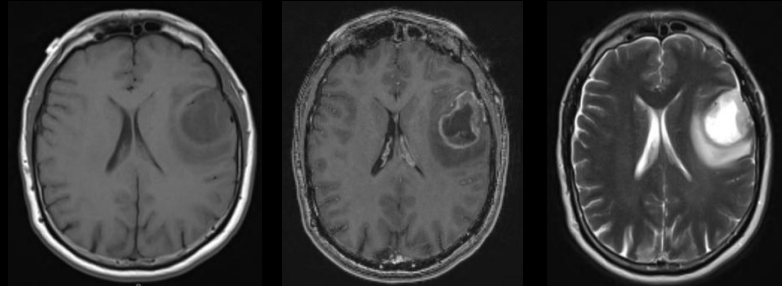
Post-Contrast
T1 Weighted
MR Image

Advanced Imaging and Analyses add to the Feature Space

Random Forests Predicted



Overlap of Radiologist Prediction-Tumor Region

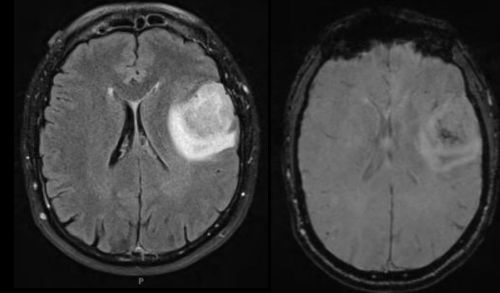


T1 pre-contrast

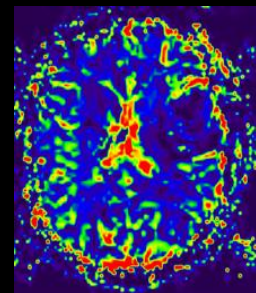
T1 post-contrast

T2

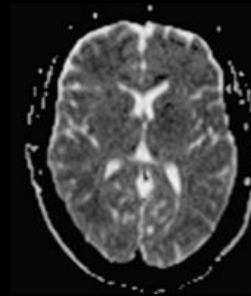
Fluid Attenuated Inversion Recovery (FLAIR)



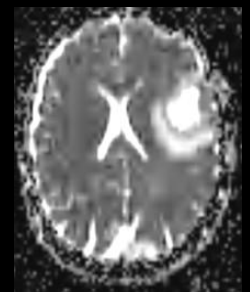
Susceptibility Weighted Imaging (SWI)



relative Cerebral Blood Volume (rCBV)



Average Diffusivity (TRACEW)



Apparent Diffusion Coefficient (ADC)

Summary

1. TCIA provides access to clinical imaging of TCGA cancers and surrounding tissues
2. Imaging data is a multi-dimensional source of quantifiable information that can complement TCGA omics
 - Spatial and temporal heterogeneity
 - Bio-functional properties
 - High dimensional feature sets
3. Ongoing research is providing promising results

Resources Links

1. Home page: <http://cancerimagingarchive.net>
2. Data sets: <https://wiki.cancerimagingarchive.net/x/mgAe>
3. TCGA Research: <https://wiki.cancerimagingarchive.net/x/sgEe>
4. Ask for more information about integrating Imaging into Omics research and analysis: cancerimagingarchive@mail.nih.gov