Decoding Breast Cancer with Quantitative Radiomics & Radiogenomics: Imaging Phenotypes in Breast Cancer Risk Assessment, Diagnosis, Prognosis, and Response to Therapy

Maryellen Giger & Yuan Ji
The University of Chicago, NorthShore University
m-giger@uchicago.edu

for the TCGA Breast Phenotype Research Group

- Analysis funded by The University of Chicago Dean Bridge Fund
- Images hosted by NCI TCIA

COI: M L Giger is a stockholder in R2/Hologic, a co-founder and equity holder in Quantitative Insights, and receives royalties from Hologic, GE Medical Systems, MEDIAN Technologies, Riverain Medical, Mitsubishi, and Toshiba

NCI TCGA/TCIA Breast Phenotype Research Group

Mapping of Breast MRI Phenotypes to Histopathology and Genomics

Computer-Extracted Phenotypes & Data analysis/associations

University of Chicago

- Maryellen Giger
- Hui Li
- Karen Drukker
- Li Lan

NorthShore University

- Yuan Ji
- Yitan Zhu
- Wentian Guo

NCI:

- Carl Jaffe
- John Freymann
- Erich Huang
- Justin Kirby
- Brenda Fevrier-Sullivan

Radiologists:

- Elizabeth Morris – MSKCC
- Ermelinda Bonaccio – Roswell
- Kathleen Brandt – Mayo
- Elizabeth Burnside – U Wisconsin Madison
- Basak Dogan – MD Anderson
- Marie Ganott – Magee
- Jose Net – U Miami
- Elizabeth Sutton – MSKCC
- Gary Whitman – MD Anderson
- Margarita Zuley – U Pittsburgh
- H. Carisa Le-Petross – MD Anderson

Human-Extracted Phenotypes Analysis

-- MD Anderson

- Arvind Rao

Decoding Breast Cancer with Quantitative Radiomics & Radiogenomics:

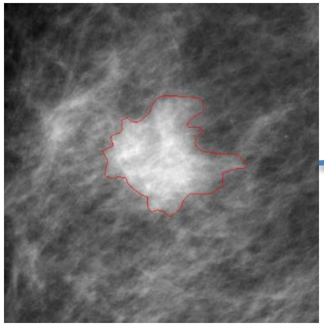
Imaging Phenotypes in Breast Cancer Risk Assessment, Diagnosis, Prognosis, and Response to Therapy

Purpose: To demonstrate, using the TCGA TCIA breast cancer dataset of MRI images, the role of quantitative radiomics in characterizing the molecular subtypes of breast cancer and associating the magenetic resonance imaging (MRI) computer-extracted image phenotypes with genomic data.

Decoding Breast Cancer with Imaging

Involves *interdisciplinary* research:

- Development and/or customization of mathematical image analysis methods for extracting **information** from biomedical image data (computer vision) - ***developed from CAD research***
- Investigations in the applications of these techniques to gain **knowledge** in (a) the ***management of the cancer patient*** and in (b) the ***understanding of cancer***



Quantitatively Extract Lesion Characteristics (Computer Vision)

Patient-Specific Image-based Tumor Signature for Precision Medicine

Data-mining of Computer-Extracted Features on Large Datasets for **Population-based** Cancer Discovery

Definitions

- Radiomics: High throughput conversion of images to mineable data

- Radiogenomics (imaging genomics): association of radiomic features with genomics and other “-omics” data

Imaging Genomics

Asks questions about the relationships between features "seen" in medical images and the biology of cancer

Data Sources

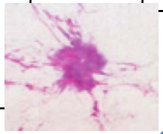
Medical Images

Computer Segmentation of Lesions

Radiologist Descriptors

Computer-extracted Lesion Features (size, morphology, texture, kinetics)

Histopathology, Molecular Classification



Genomics

Associations and/or Classification Relevant to Clinical or Biological Questions – Develop Predictive Models

Imaging Genomics

Asks questions about the relationships between features "seen" in medical images and the biology of cancer

Data Sources

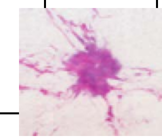
Medical Images

Computer Segmentation of Lesions

Computer-extracted Lesion Features (size, morphology, texture, kinetics)

Radiologist Descriptors

Histopathology, Molecular Classification



Genomics

Which correlate and which are complementary???

Associations and/or Classification Relevant to Clinical or Biological Questions – Develop Predictive Models

Imaging Genomics

Asks questions about the relationships between features "seen" in medical images and the biology of cancer

Data Sources

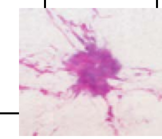
Medical Images

Computer Segmentation of Lesions

Radiologist Descriptors

Computer-extracted Lesion Features (size, morphology, texture, kinetics)

Histopathology, Molecular Classification



Genomics

Lead to Personalized Screening and Personalized Treatment

Which correlate and which are complementary???

Associations and/or Classification Relevant to Clinical or Biological Questions – Develop Predictive Models

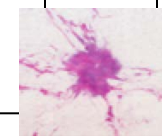
Imaging Genomics

Asks questions about the relationships between features "seen" in medical images and the biology of cancer

Data Sources

Medical Images

Histopathology, Molecular Classification



Genomics

Radiomics

Computer Segmentation of Lesions

Computer-extracted Lesion Features (size, morphology, texture, kinetics)

Radiologist Descriptors

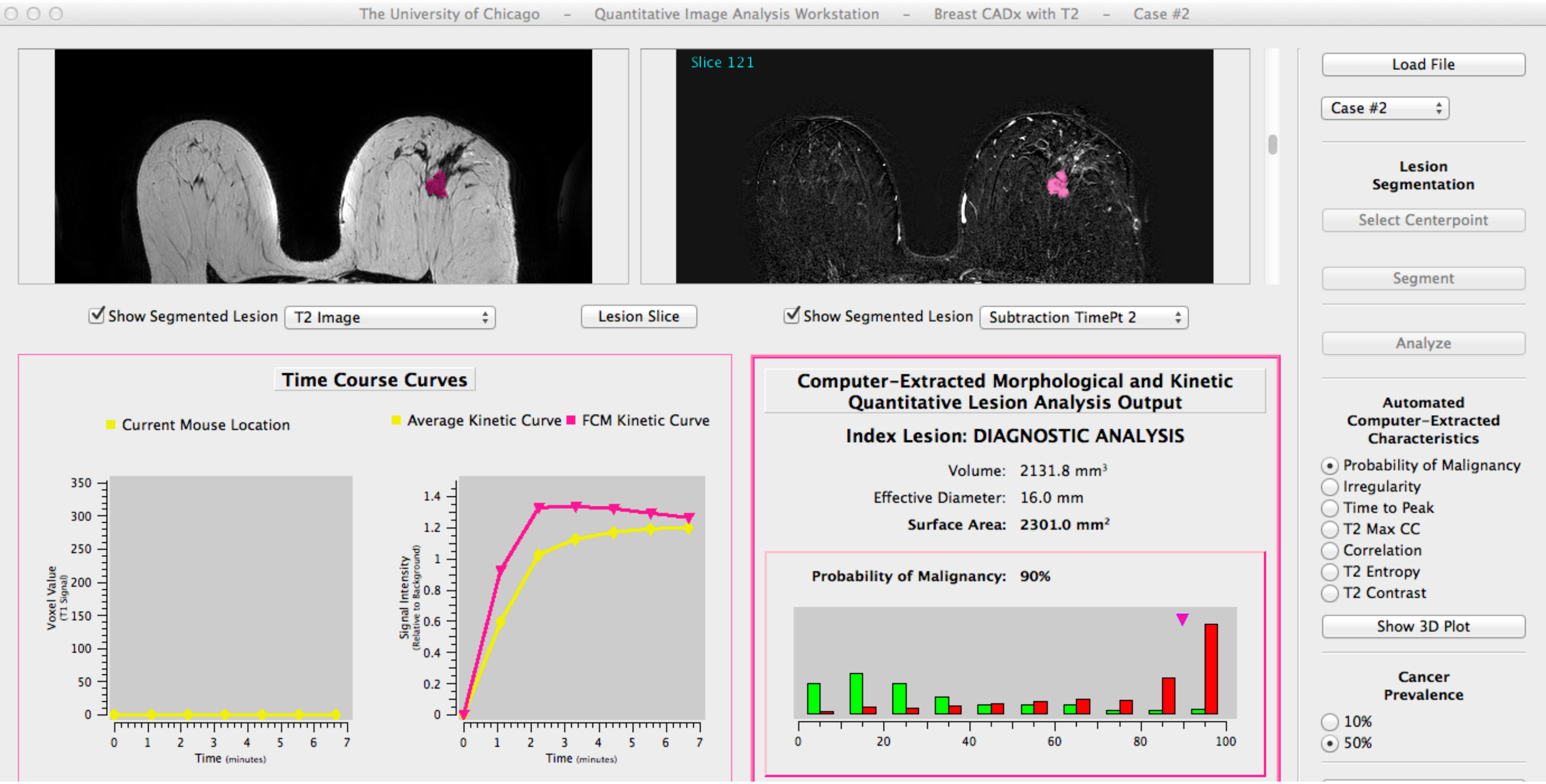
Lead to Personalized Screening and Personalized Treatment

Which correlate and which are complementary???

Associations and/or Classification Relevant to Clinical or Biological Questions – Develop Predictive Models

University of Chicago High-Throughput MRI Phenotyping System (Quantitative Image Analysis Workstation)

Automated Lesion Segmentation, Feature Extraction [volumetrics, morphological, texture, kinetics] and Estimation of the Probability of Malignancy



Giger et al., RSNA 2010

Dataset

The Cancer Genome Atlas

cancergenome.nih.gov

THE CANCER IMAGING ARCHIVE

cancerimagingarchive.net

Breast Cancer cases

Clinical /Histopathology /Genomic data downloaded by TCGA Assembler & Molecular subtyping / risk of recurrence values by Perou Lab

MRIs of 91 cases (GE 1.5T) collected by TCIA

Tumor location on MRI determined by consensus of three of the TCIA radiologists

MRIs of 91 cases downloaded to UChicago for computational MRI tumor phenotyping (radiomics)

Dataset

The Cancer Genome Atlas

cancergenome.nih.gov

THE CANCER IMAGING ARCHIVE

cancerimagingarchive.net

Breast Cancer cases

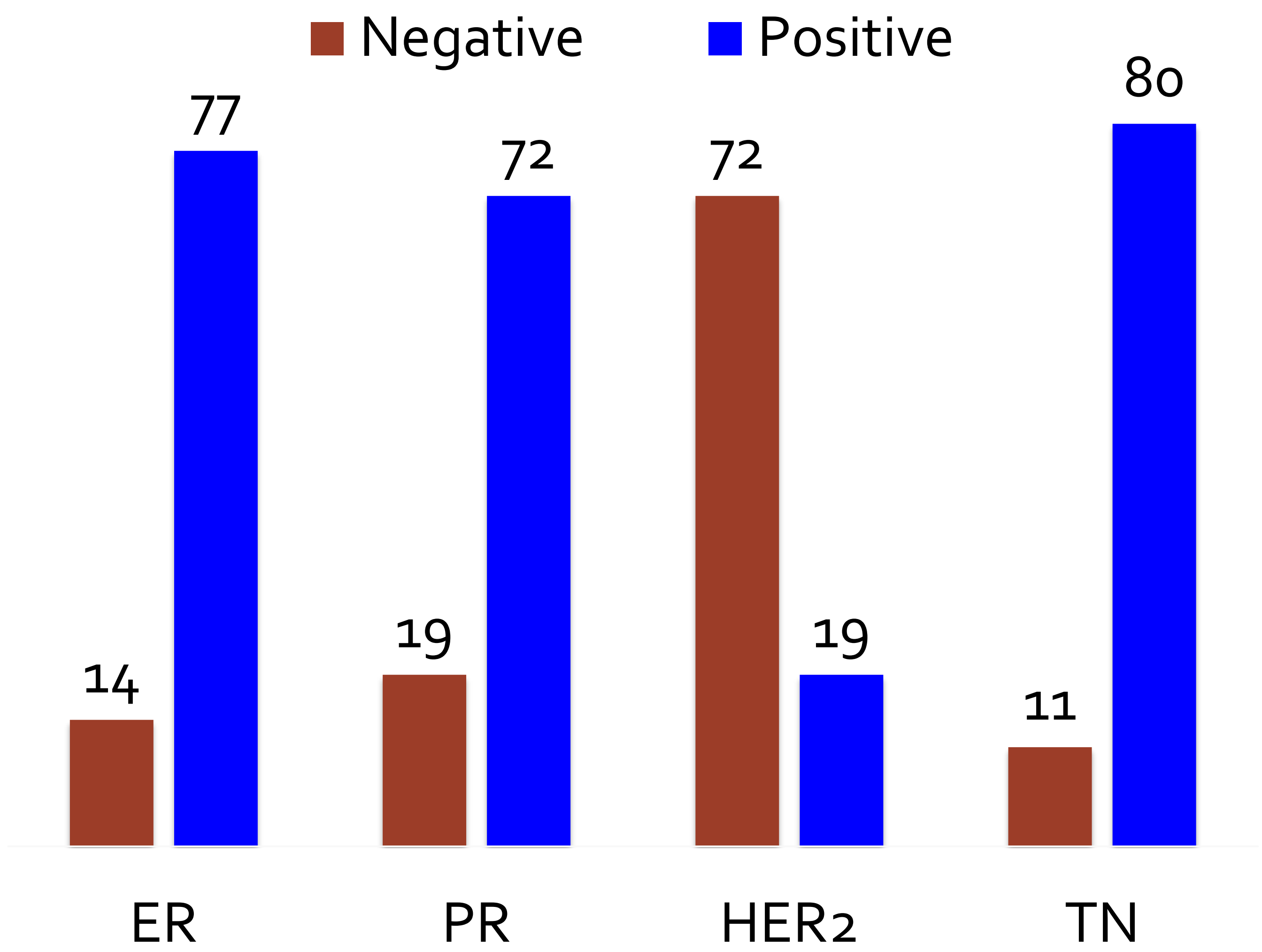
Clinical /Histopathology /Genomic data downloaded by TCGA Assembler & Molecular subtyping / risk of recurrence values by Perou Lab

MRIs of 91 cases (GE 1.5T) collected by TCIA

Tumor location on MRI determined by consensus of three of the TCIA radiologists

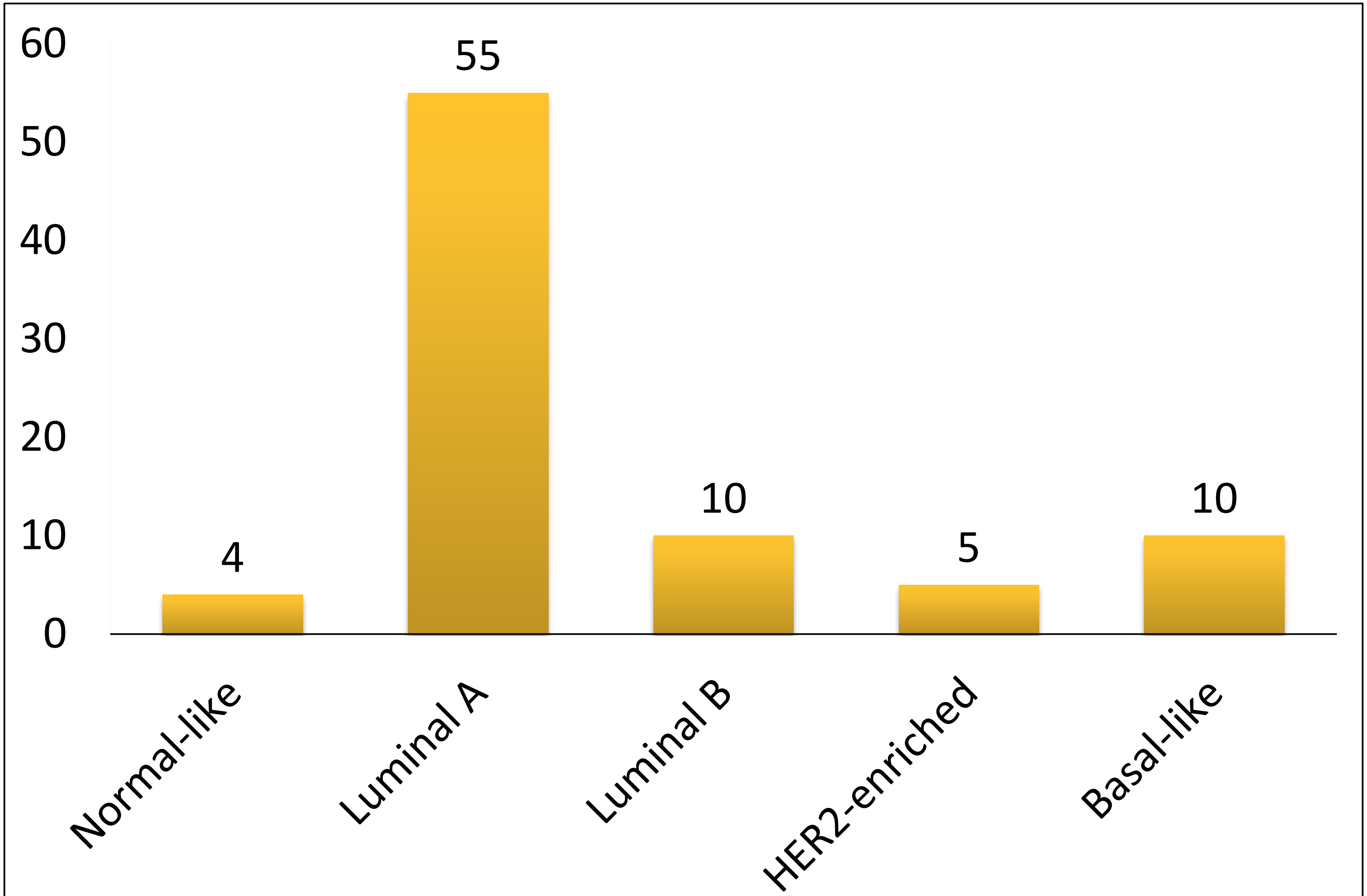
MRIs of 91 cases downloaded to UChicago for computational MRI tumor phenotyping (radiomics)

Distribution of the 91 MRI cases



Giger TCGA 2015

Distribution of the 91 MRI cases



Giger TCGA 2015

Dataset

The Cancer Genome Atlas

cancergenome.nih.gov

THE CANCER IMAGING ARCHIVE

cancerimagingarchive.net

Breast Cancer cases

Clinical /Histopathology /Genomic data downloaded by TCGA Assembler & Molecular subtyping / risk of recurrence values by Perou Lab

MRIs of 91 cases (GE 1.5T) collected by TCIA

Tumor location on MRI determined by consensus of three of the TCIA radiologists

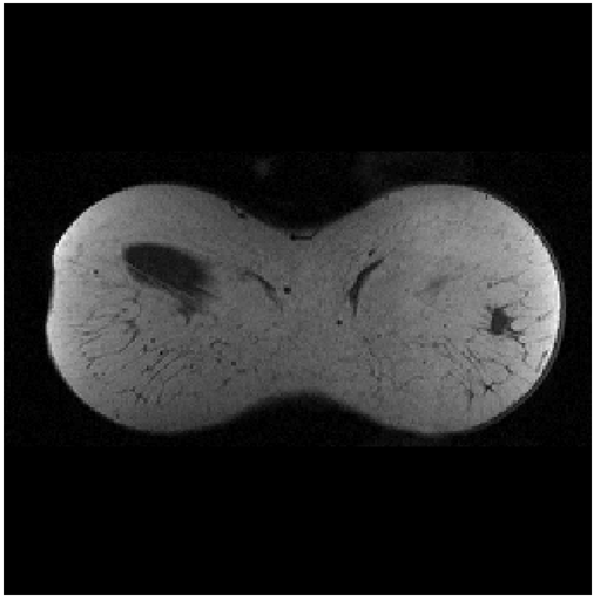
MRIs of 91 cases downloaded to UChicago for computational MRI tumor phenotyping (radiomics)

Contrast-enhanced MR images of breast

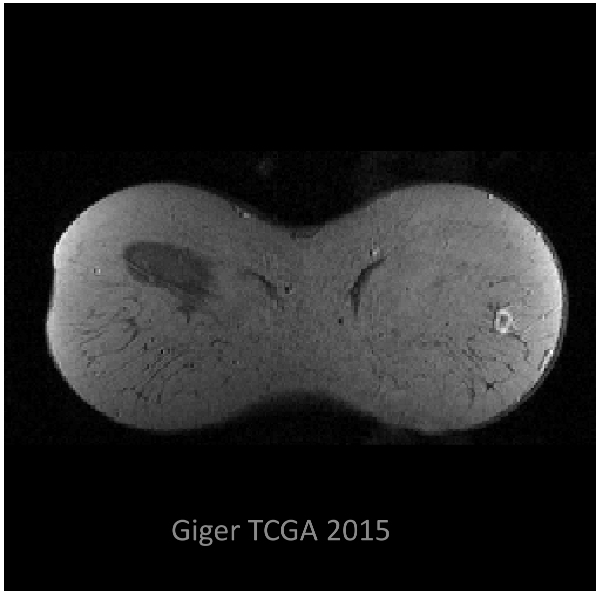
- **Tumors have increased blood vessels and differ in microvascular density and vessel permeability**

- **Gd-DTPA shortens T1 relaxation time which leads to increase of signal in T1-weighted images**

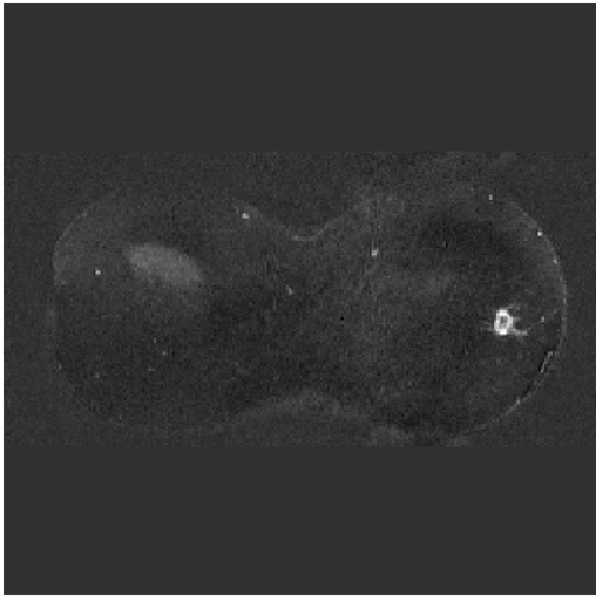
Precontrast



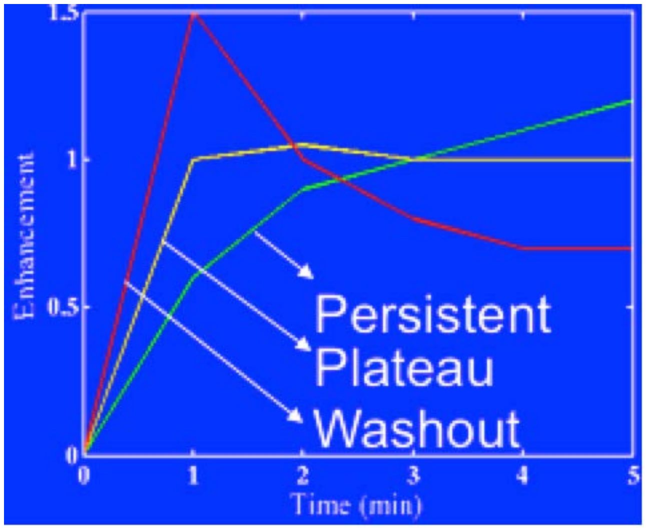
Postcontrast



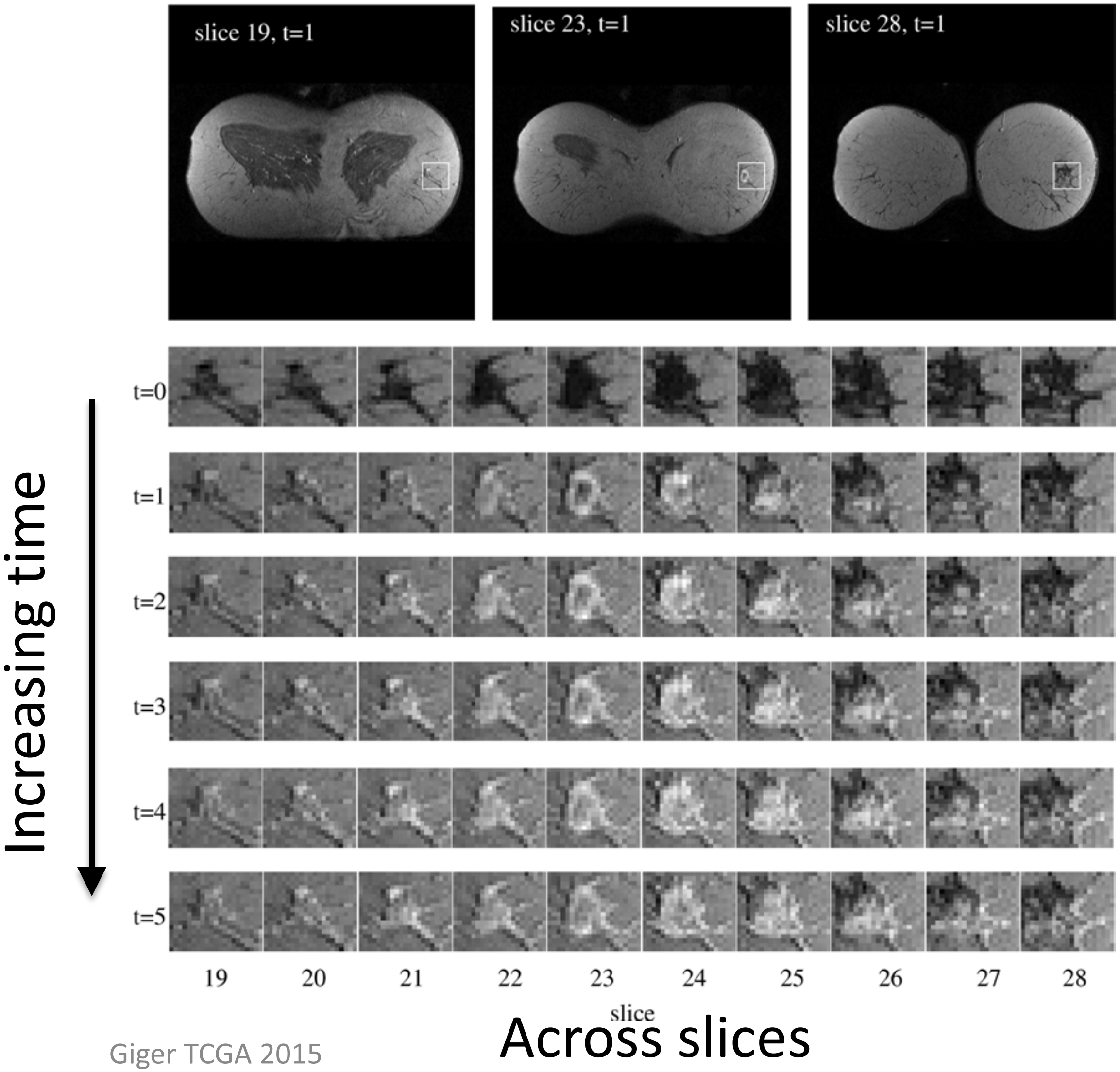
Subtraction



Dynamic Contrast-Enhanced MRI & Tumor Segmentation



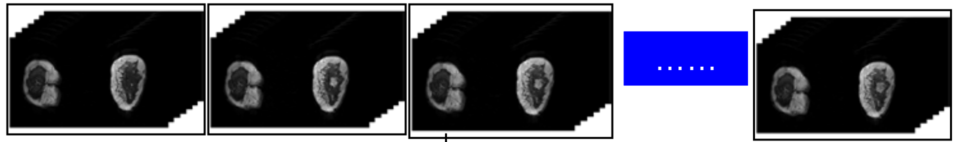
4D image analysis



Giger TCGA 2015

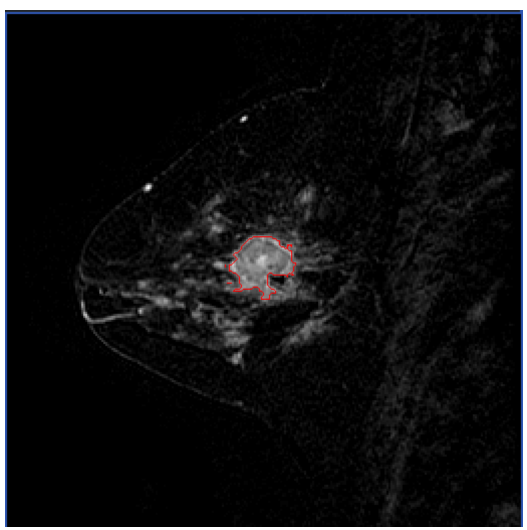
University of Chicago High-Throughput MRI Phenotyping System (Segmentation of the Tumor within the Breast MR image)

4D DCE MRI images

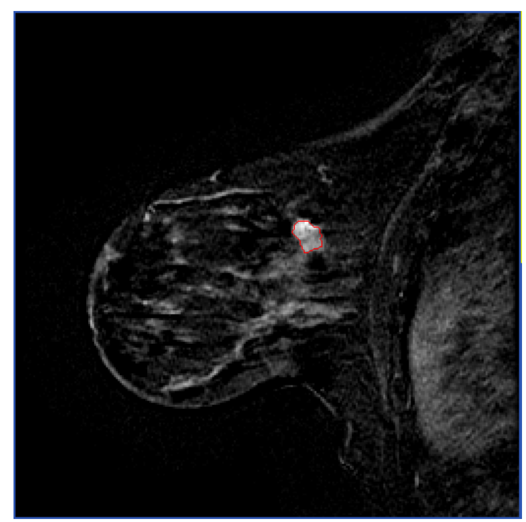


Radiologist-indicated Tumor Center

Computerized Tumor Segmentation

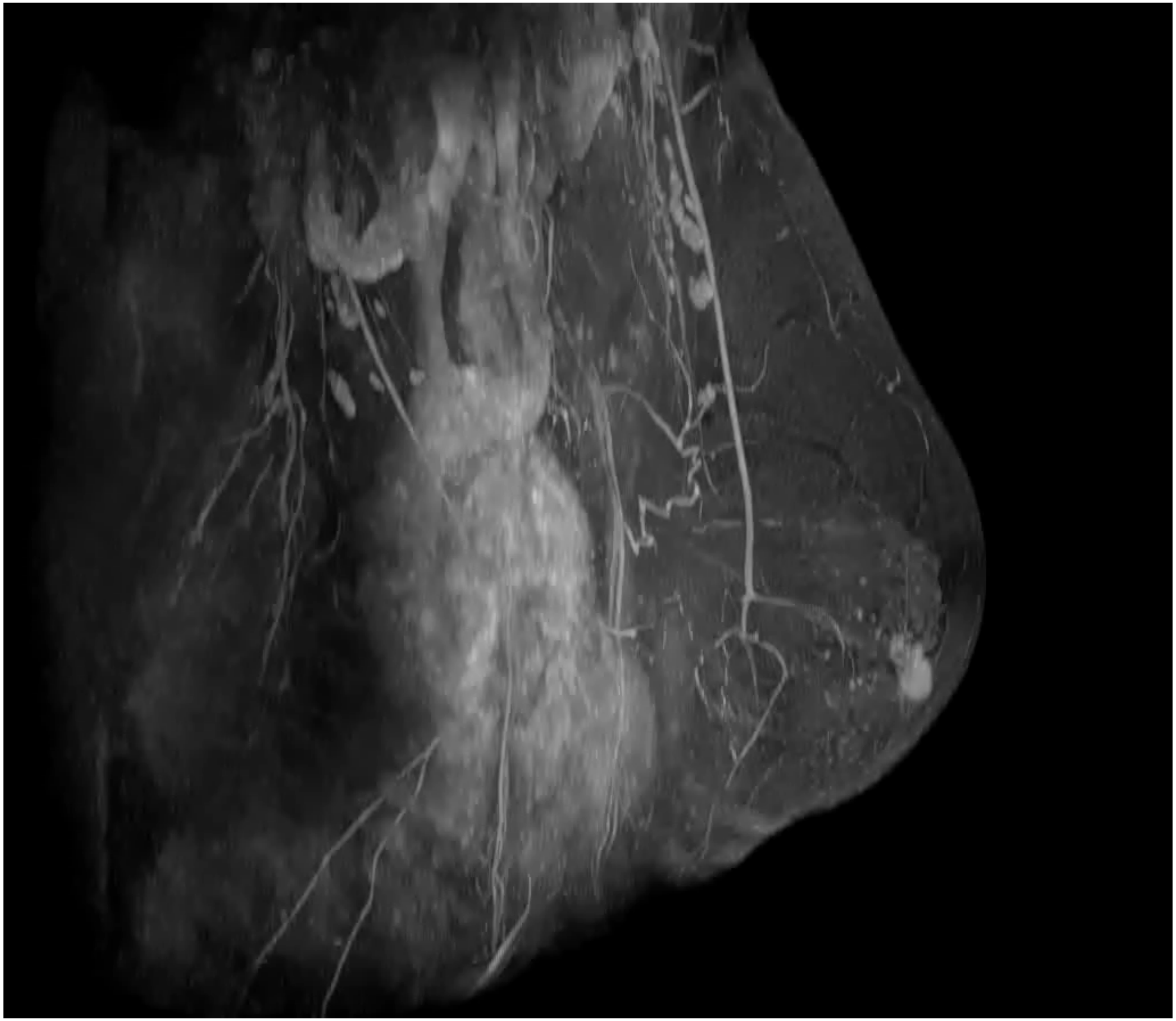


ER-negative

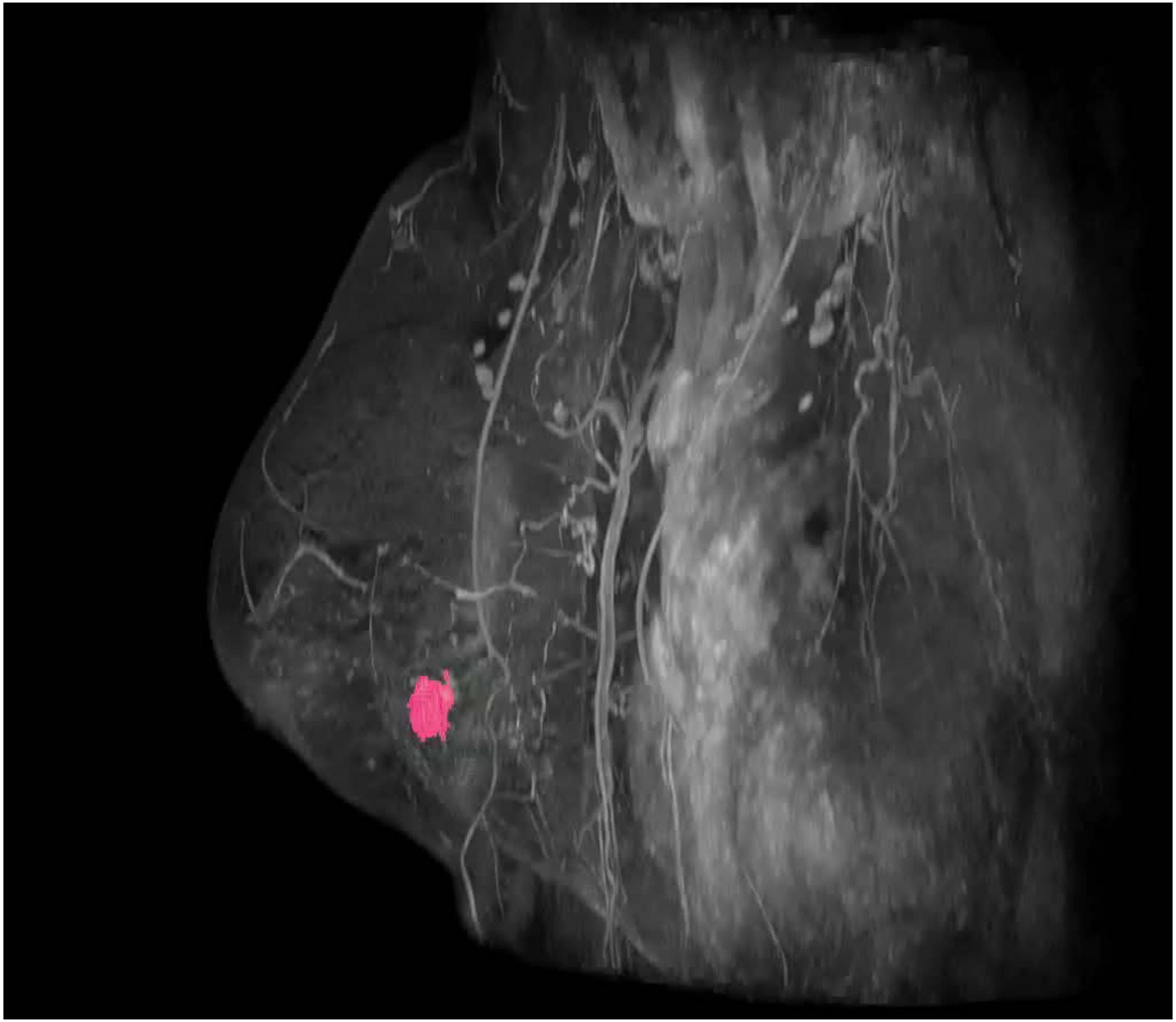


ER-positive

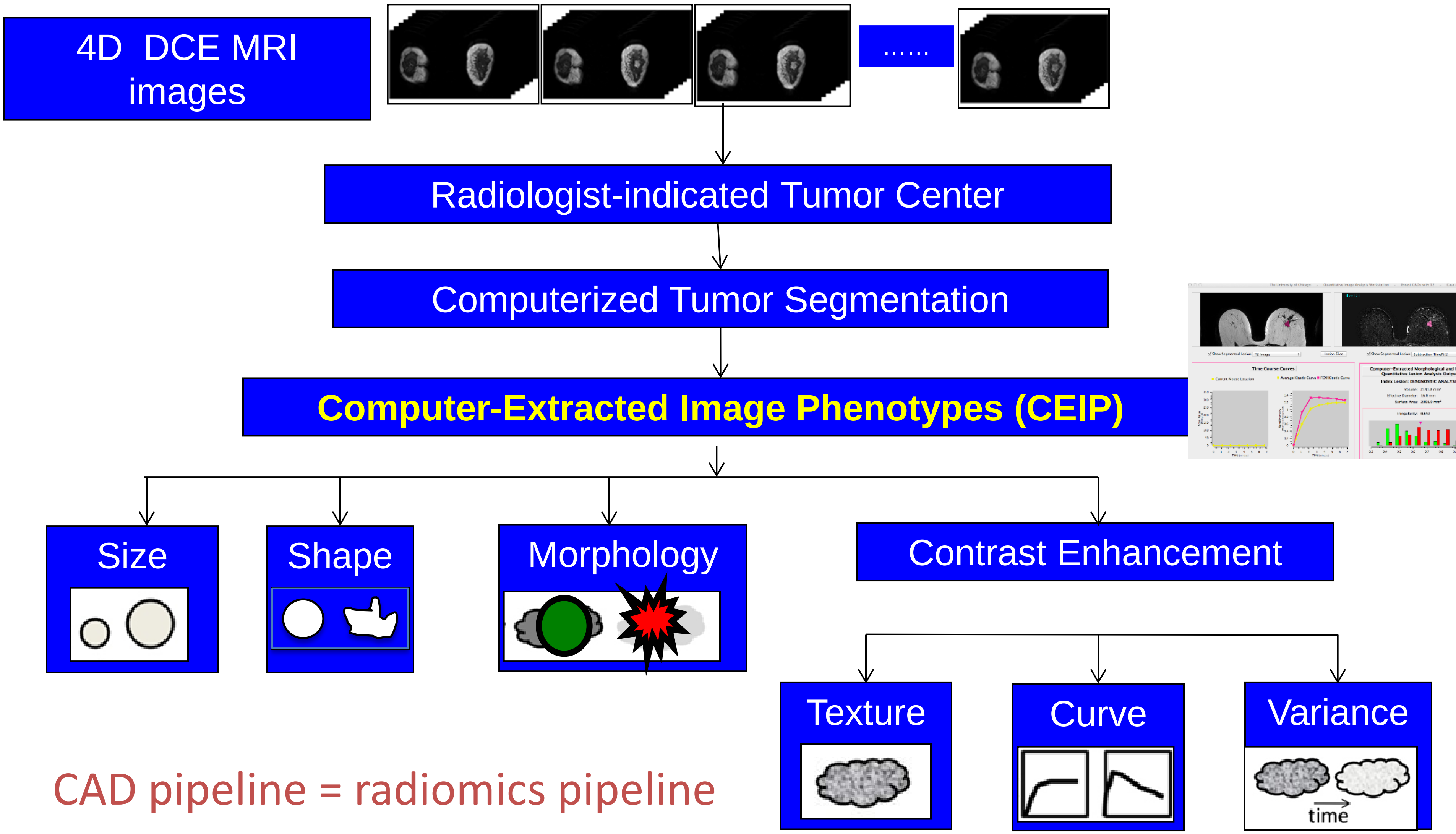
3D Breast MRI image



Computer-extracted Breast Cancer on MRI (can analyze as a "virtual" biopsy of the tumor)

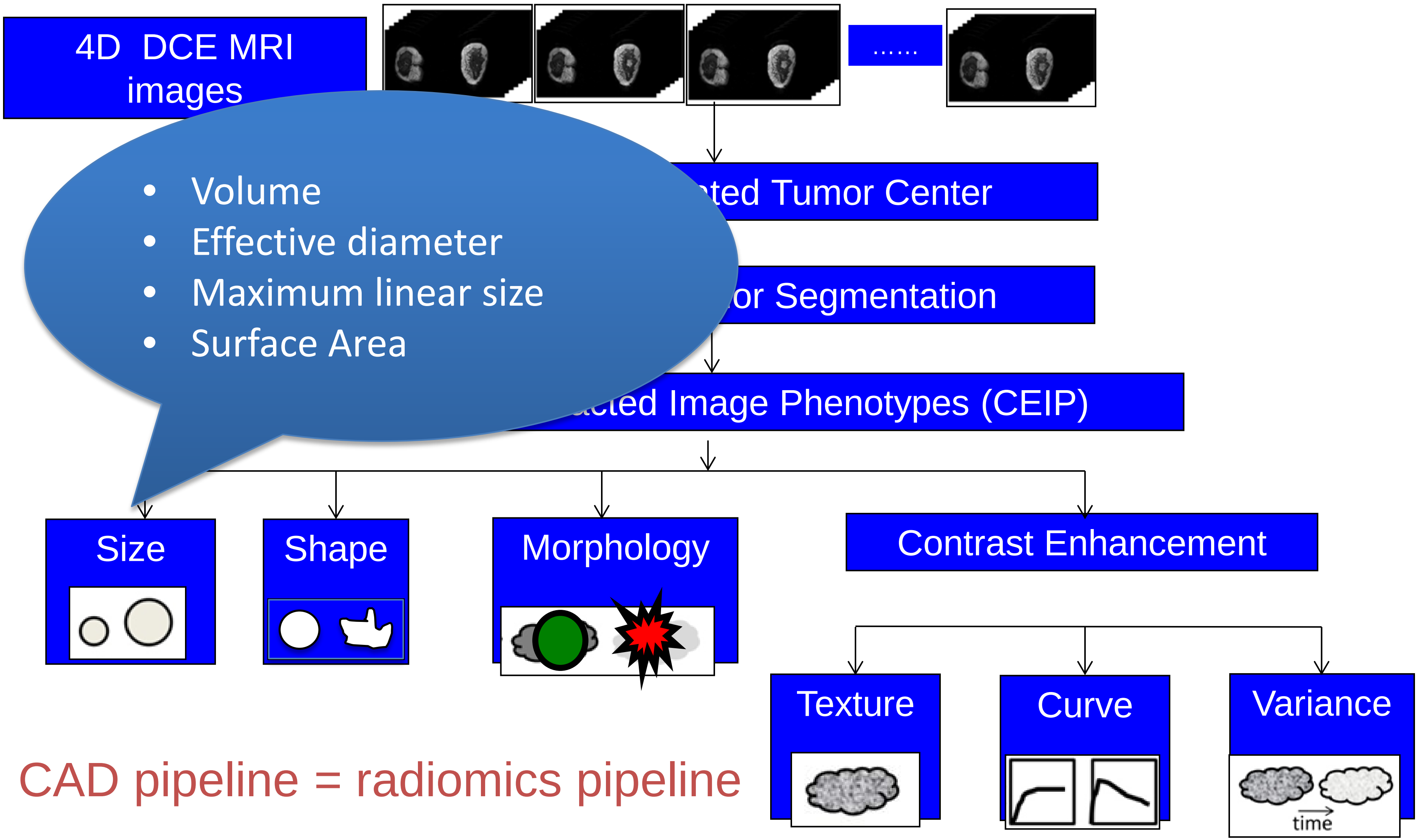


University of Chicago High-Throughput MRI Phenotyping System



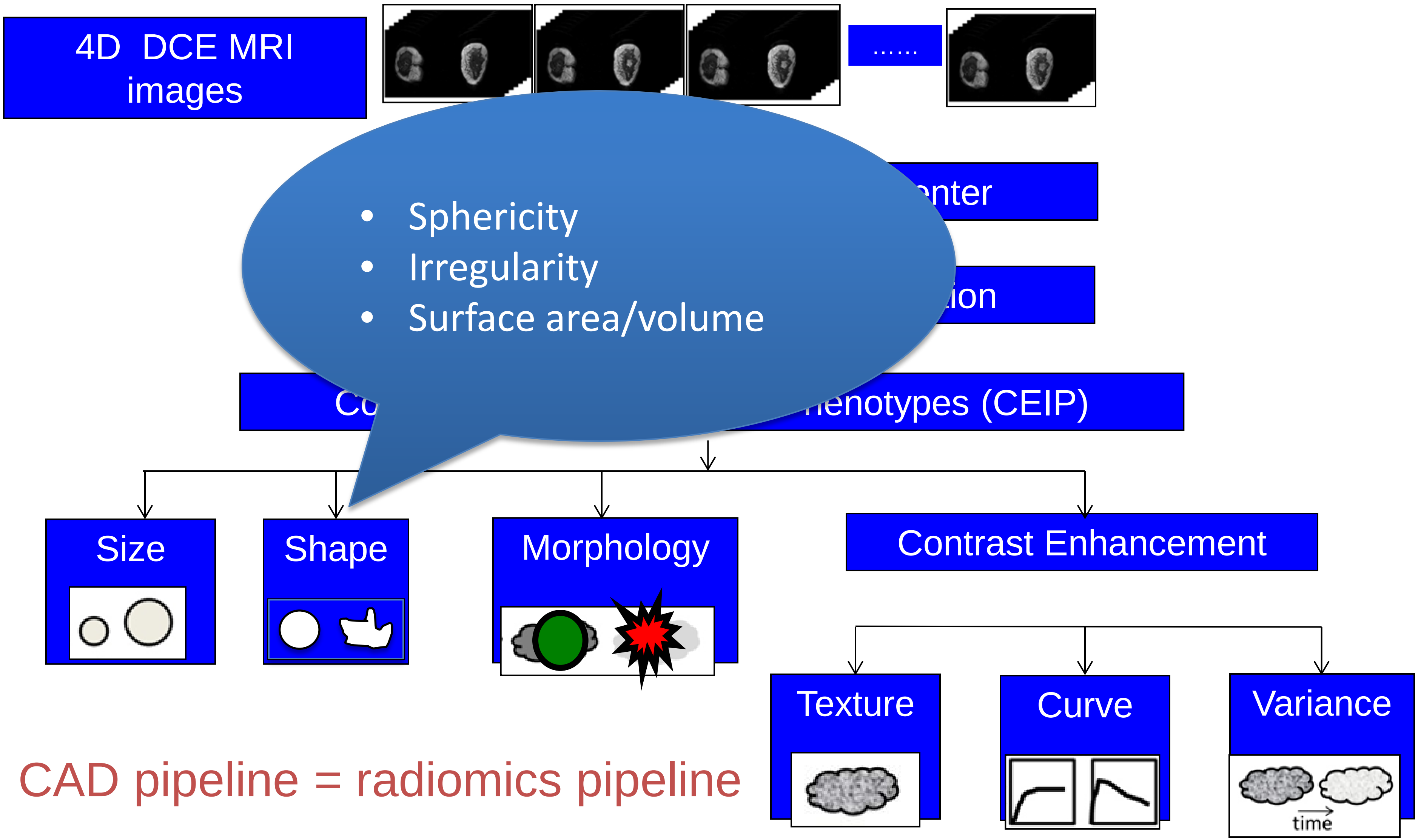
CAD pipeline = radiomics pipeline

University of Chicago High-Throughput MRI Phenotyping System



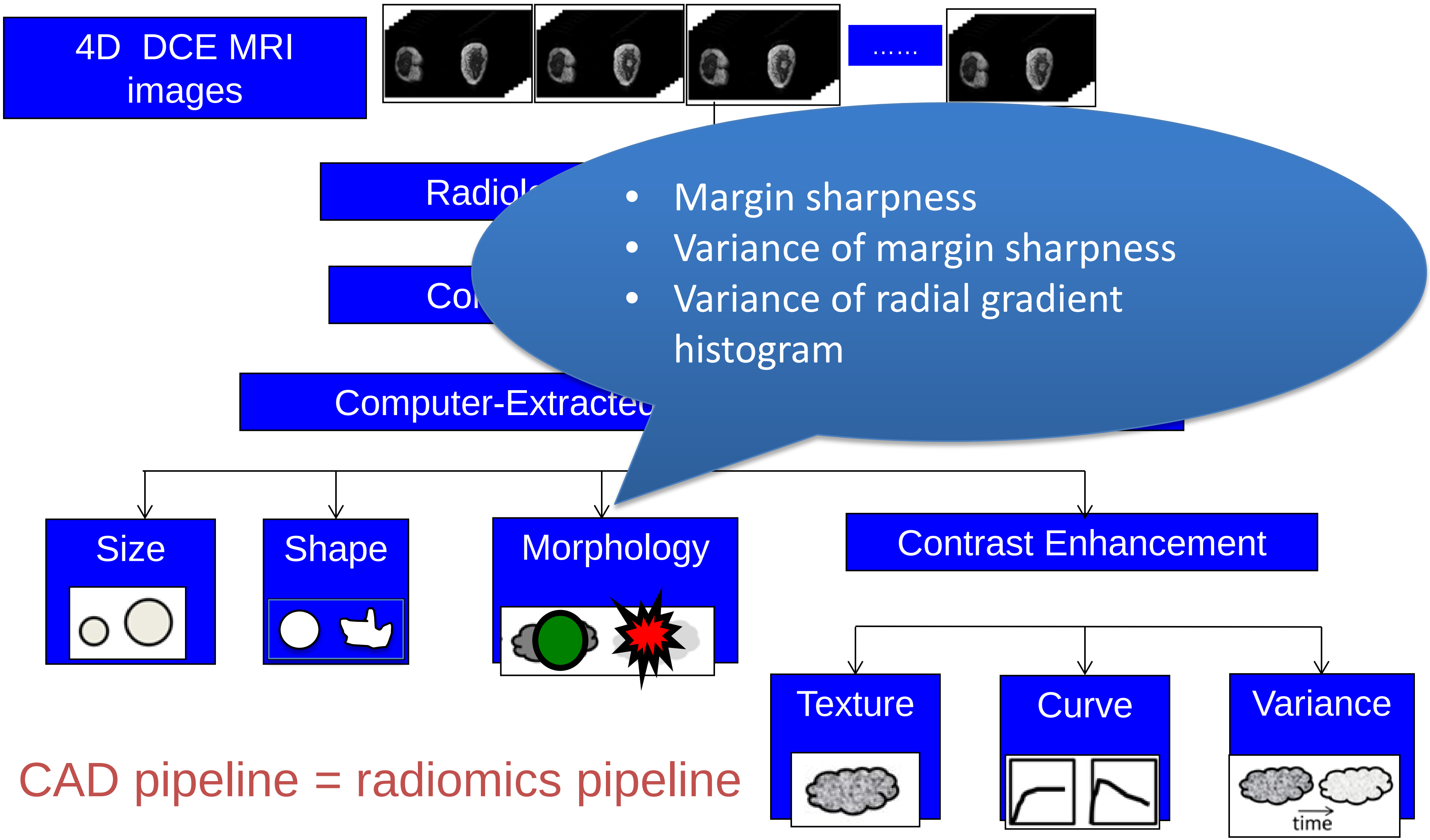
CAD pipeline = radiomics pipeline

University of Chicago High-Throughput MRI Phenotyping System



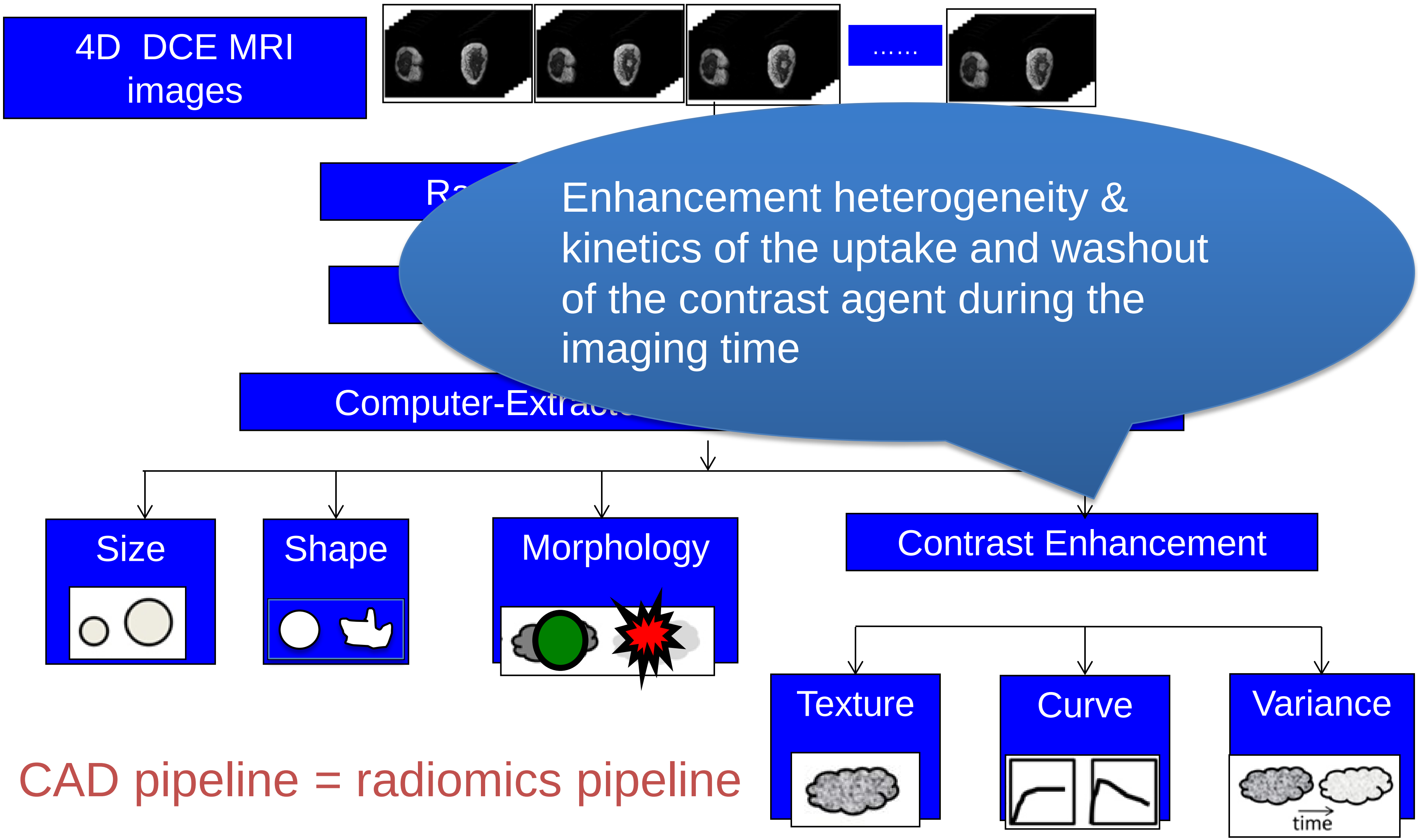
CAD pipeline = radiomics pipeline

University of Chicago High-Throughput MRI Phenotyping System



CAD pipeline = radiomics pipeline

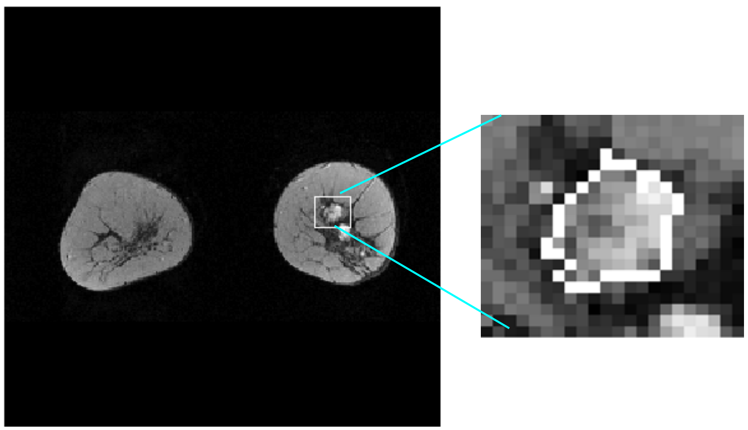
University of Chicago High-Throughput MRI Phenotyping System

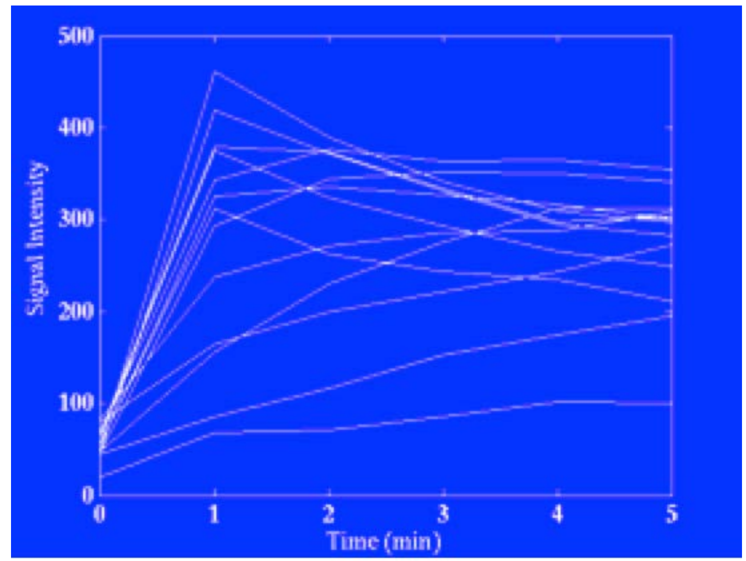


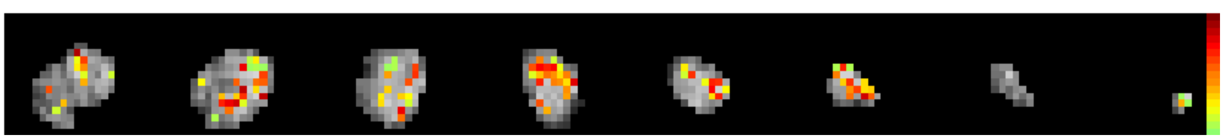
CAD pipeline = radiomics pipeline

Tumors are Heterogeneous:
Contrast Enhancement Heterogeneity & Kinetics

Heterogeneity of Tumors:

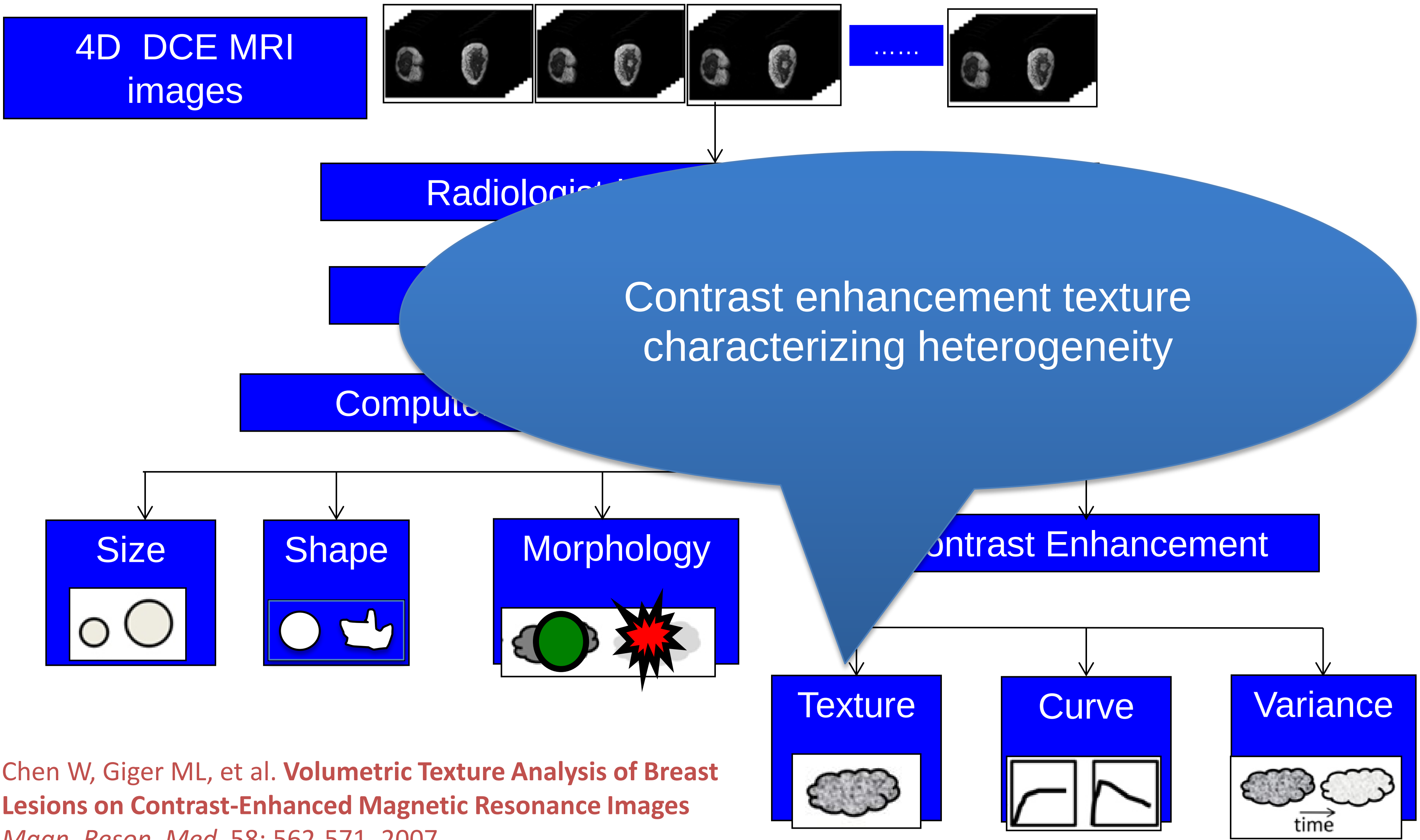






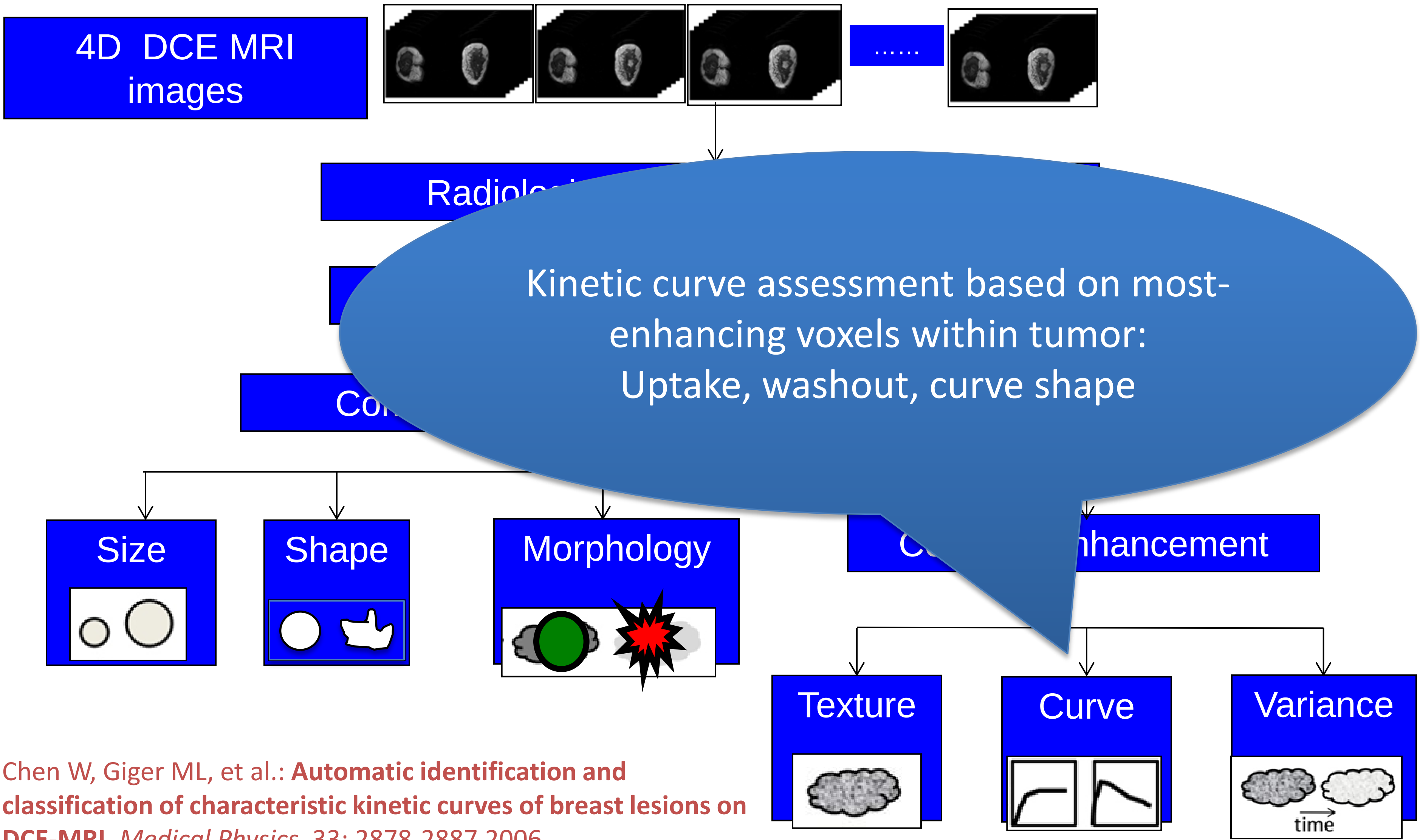
Giger TCGA 2015

University of Chicago High-Throughput MRI Phenotyping System



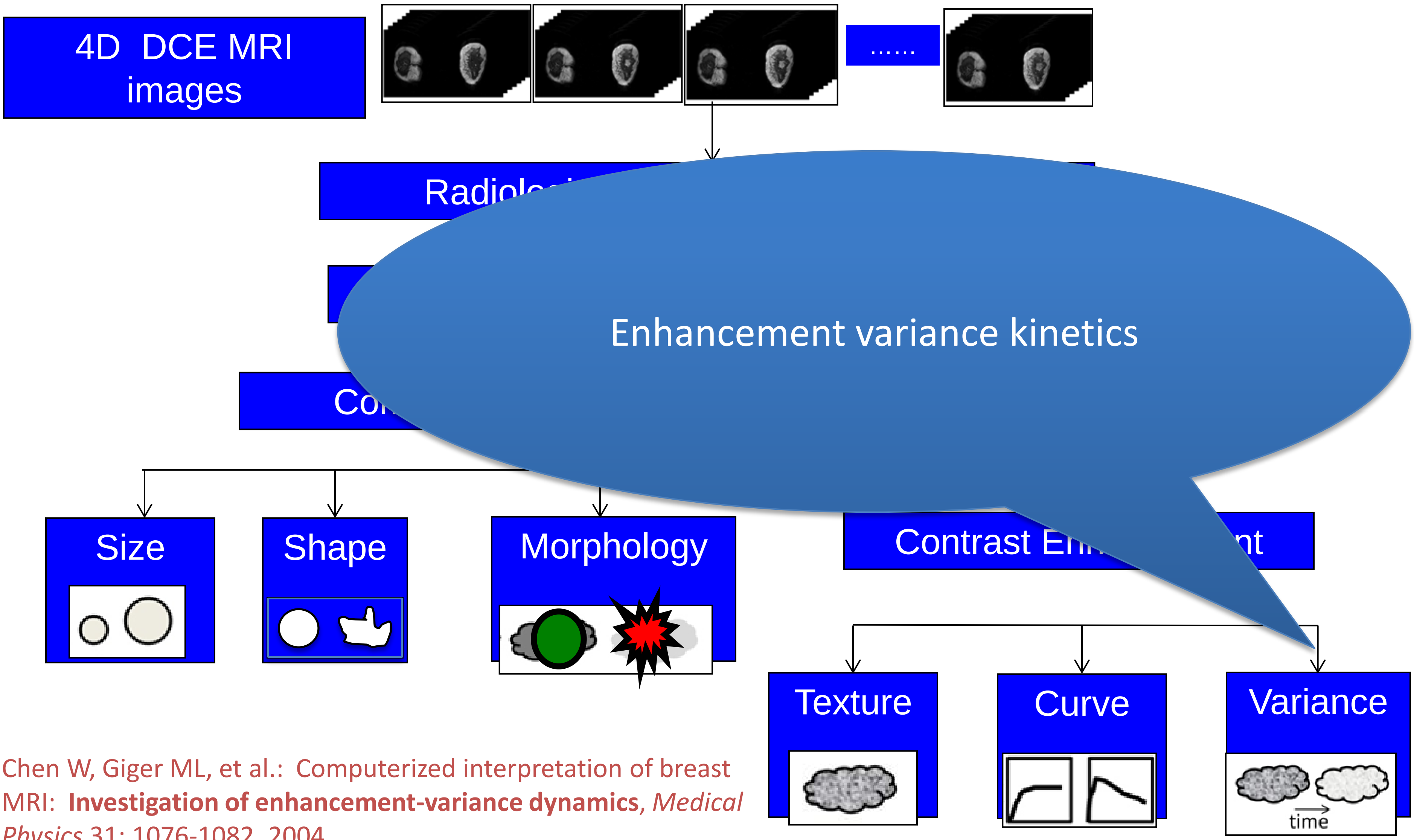
Chen W, Giger ML, et al. **Volumetric Texture Analysis of Breast Lesions on Contrast-Enhanced Magnetic Resonance Images** *Magn. Reson. Med.* 58: 562-571, 2007

University of Chicago High-Throughput MRI Phenotyping System



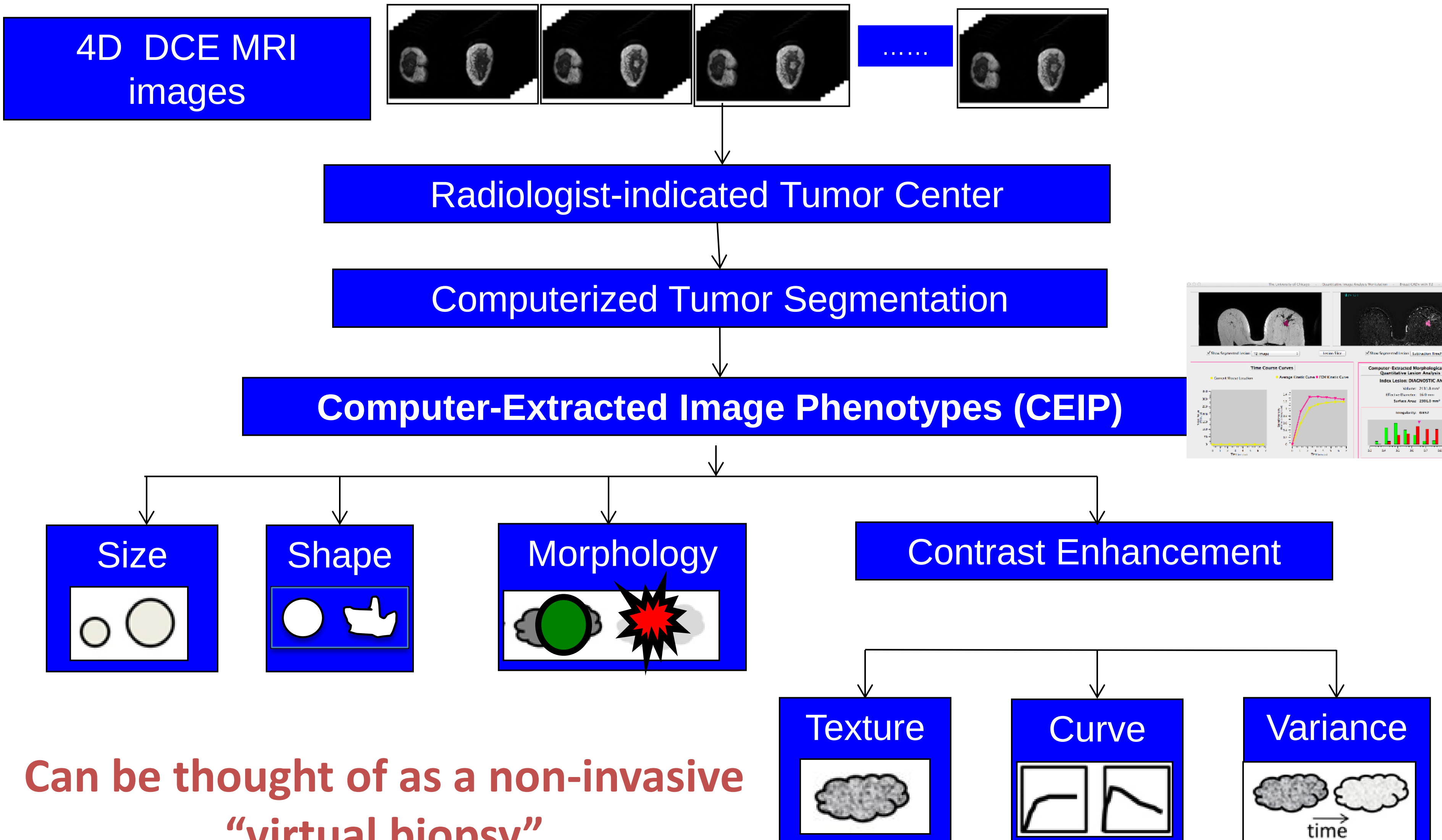
Chen W, Giger ML, et al.: **Automatic identification and classification of characteristic kinetic curves of breast lesions on DCE-MRI.** *Medical Physics*, 33: 2878-2887,2006

University of Chicago High-Throughput MRI Phenotyping System



Chen W, Giger ML, et al.: Computerized interpretation of breast MRI: **Investigation of enhancement-variance dynamics**, *Medical Physics* 31: 1076-1082, 2004

University of Chicago High-Throughput MRI Phenotyping System For Breast Tumors



Can be thought of as a non-invasive "virtual biopsy"

Giger TCGA 2015

“Virtual biopsy” yielding tumor phenotypes & signatures

Relating Computer-extracted MRI Phenotypes to:

Classification & Association Tasks:

1. Clinical Tumor Status
    1. Tumor Stage
    2. Presence or Absence of Positive Lymph Nodes
2. Molecular Classification & Cancer Subtype
    1. ER- vs. ER+
    2. PR- vs. PR+
    3. Her2- vs. Her2+
    4. Triple Negative vs. Others
3. Risk of Recurrence
    1. OncotypeDX
    2. PAM50
    3. MammaPrint
4. Genomic Pathways

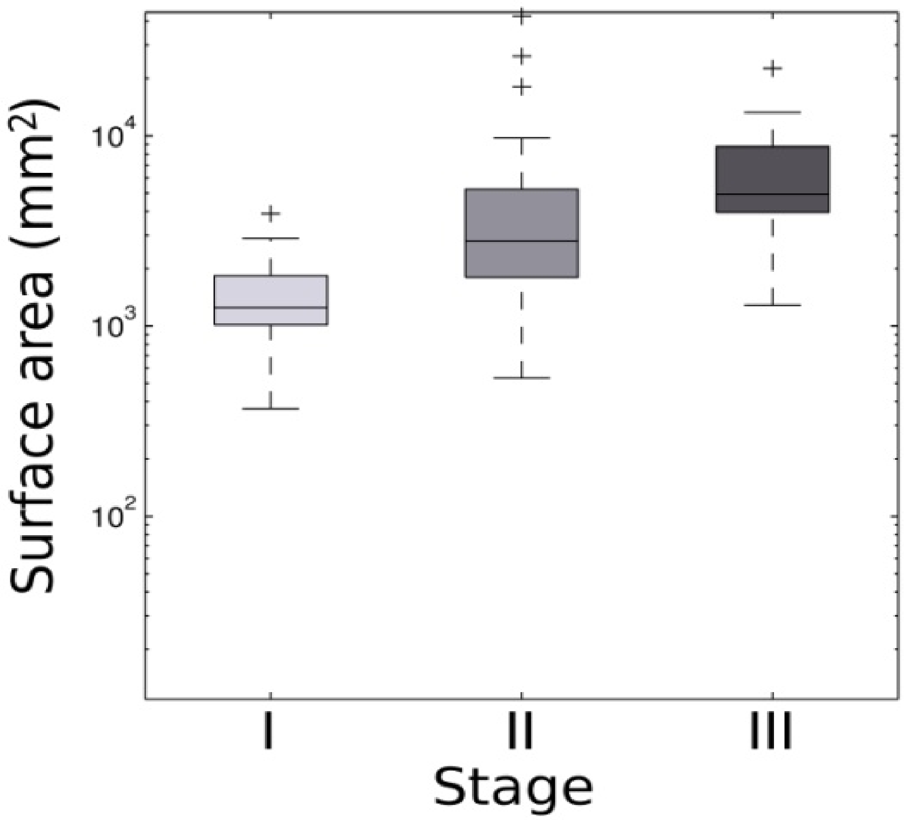
“Virtual biopsy” yielding tumor phenotypes & signatures

Relating Computer-extracted MRI Phenotypes to:

Classification & Association Tasks:

1. Clinical Tumor Status
    1. Tumor Stage
    2. Presence or Absence of Positive Lymph Nodes
2. Molecular Classification & Cancer Subtype
    1. ER- vs. ER+
    2. PR- vs. PR+
    3. Her2- vs. Her2+
    4. Triple Negative vs. Others
3. Risk of Recurrence & Cancer Subtype (Normal, Luminal A..)
    1. OncotypeDX
    2. PAM50
    3. MammaPrint
4. Genomic Pathways

MRI-based Phenotypes of Size – predictive of breast cancer tumor stage



TCGA/TCIA Breast Cancer Group cases;
University of Chicago Giger Lab computer-extracted image phenotypes

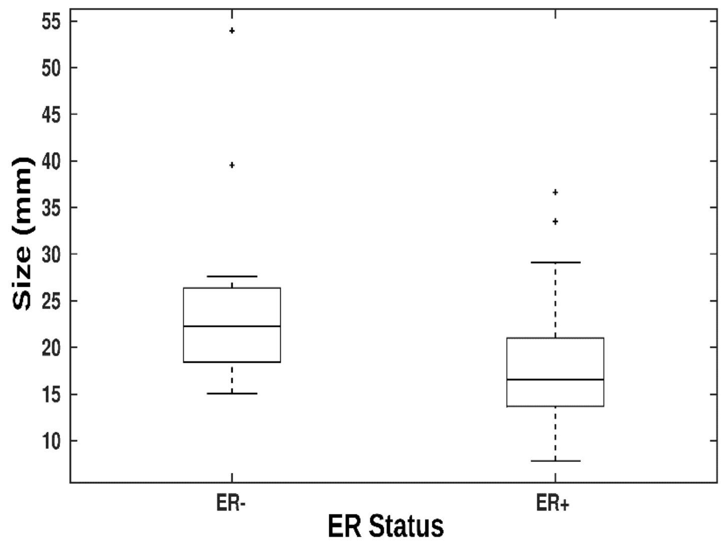
“Virtual biopsy” yielding tumor phenotypes & signatures

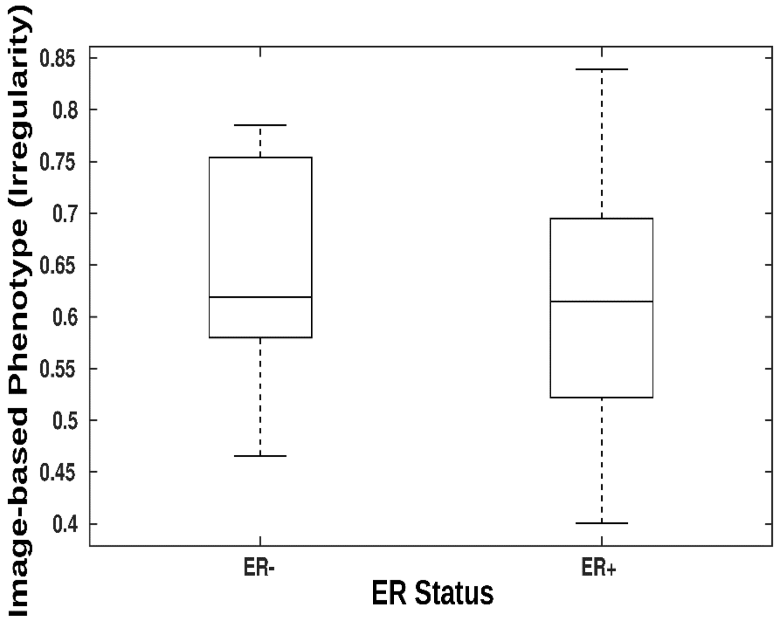
Relating Computer-extracted MRI Phenotypes to:

Classification & Association Tasks:

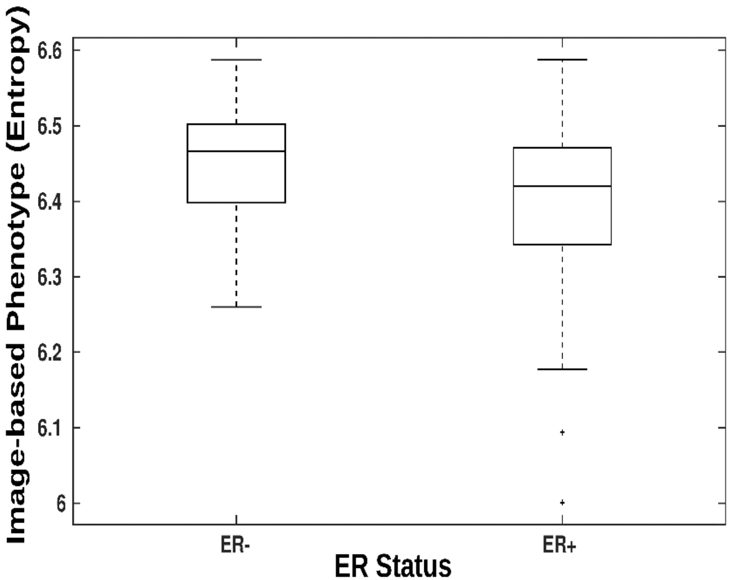
1. Clinical Tumor Status
    1. Tumor Stage
    2. Presence or Absence of Positive Lymph Nodes
2. Molecular Classification & Cancer Subtype
    1. ER- vs. ER+
    2. PR- vs. PR+
    3. HER2- vs. HER2+
    4. Triple Negative vs. Others
3. Risk of Recurrence
    1. OncotypeDX
    2. PAM50
    3. MammaPrint
4. Genomic Pathways

From TCIA MRI Radiomics -- ER Negative Breast Cancers tended to have larger size, a more irregular shape, and more heterogeneous in terms of contrast enhancement

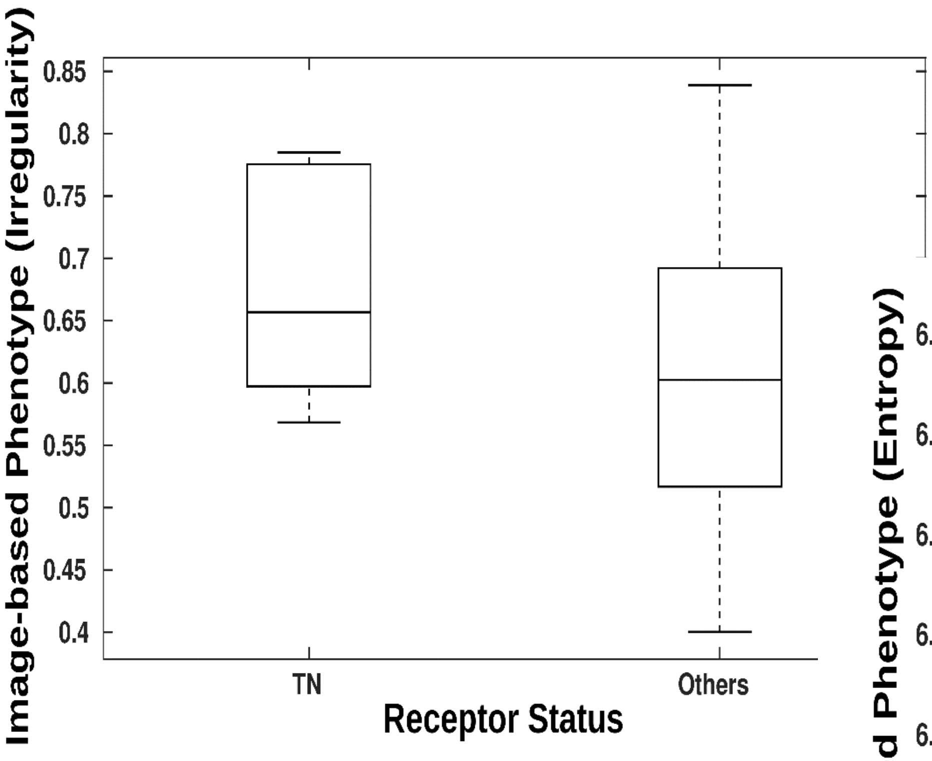


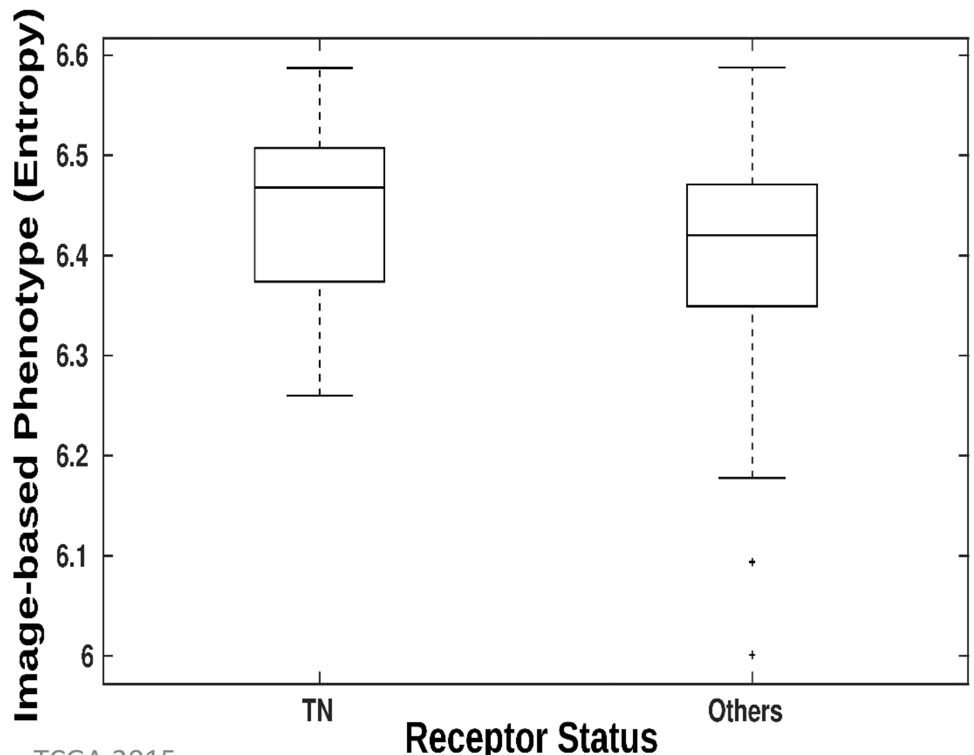


TCGA 2015



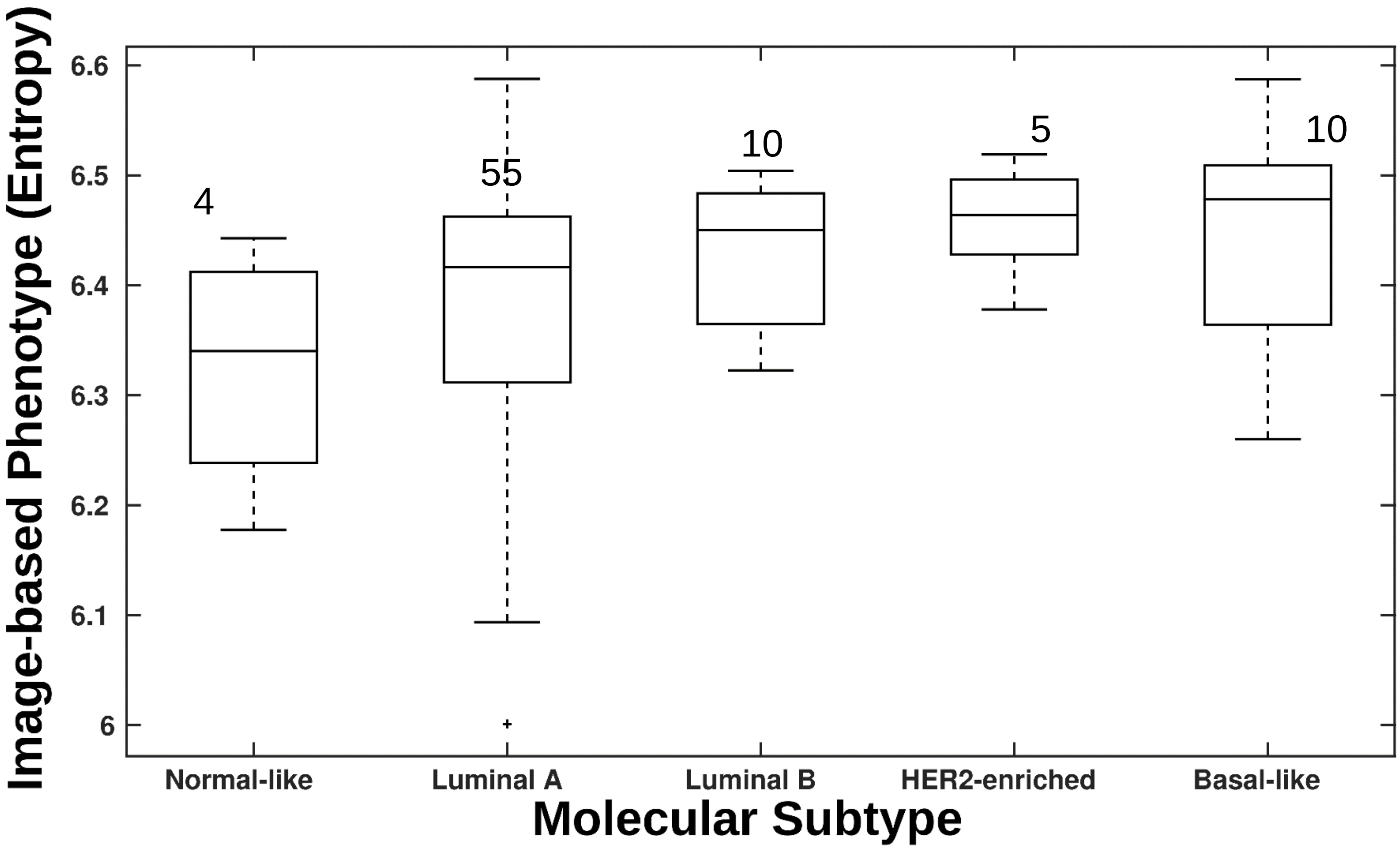
From TCIA Radiomics- Triple Negative Breast Cancers tended to have a more irregular shape, and more heterogeneous in terms of contrast enhancement





Giger TCGA 2015

From the TCIA Radiomics -- Enhancement Texture of Tumor Heterogeneity appears Predictive of Molecular Subtype

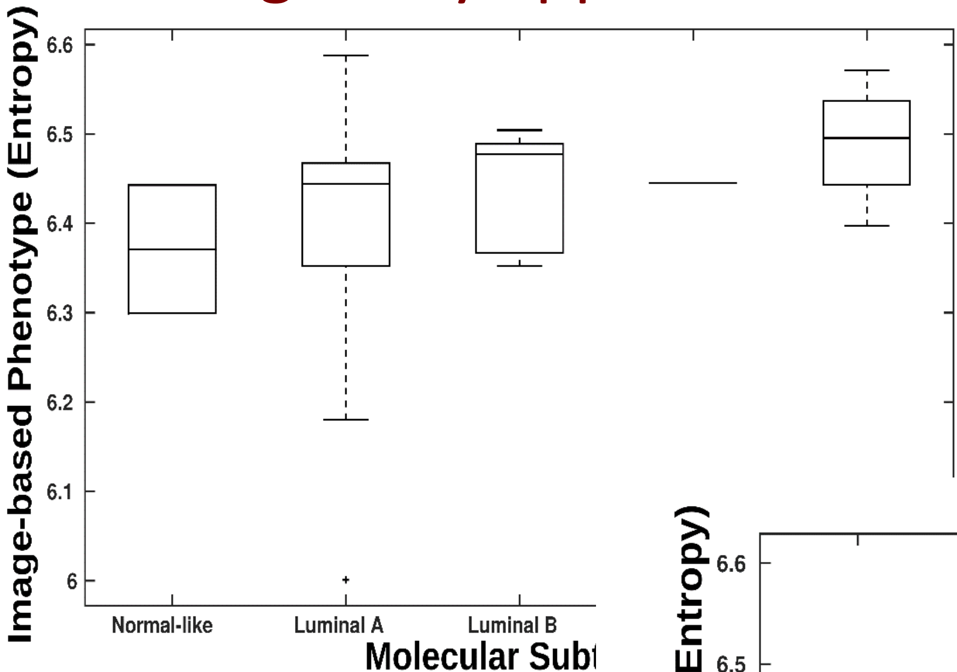


Kendall test results for trends; p-value=0.0055

Giger TCGA 2015

Molecular Subtyping from C. Perou

From the TCIA Radiomics -- Enhancement Texture of Tumor Heterogeneity appears **Predictive of Molecular Subtype**

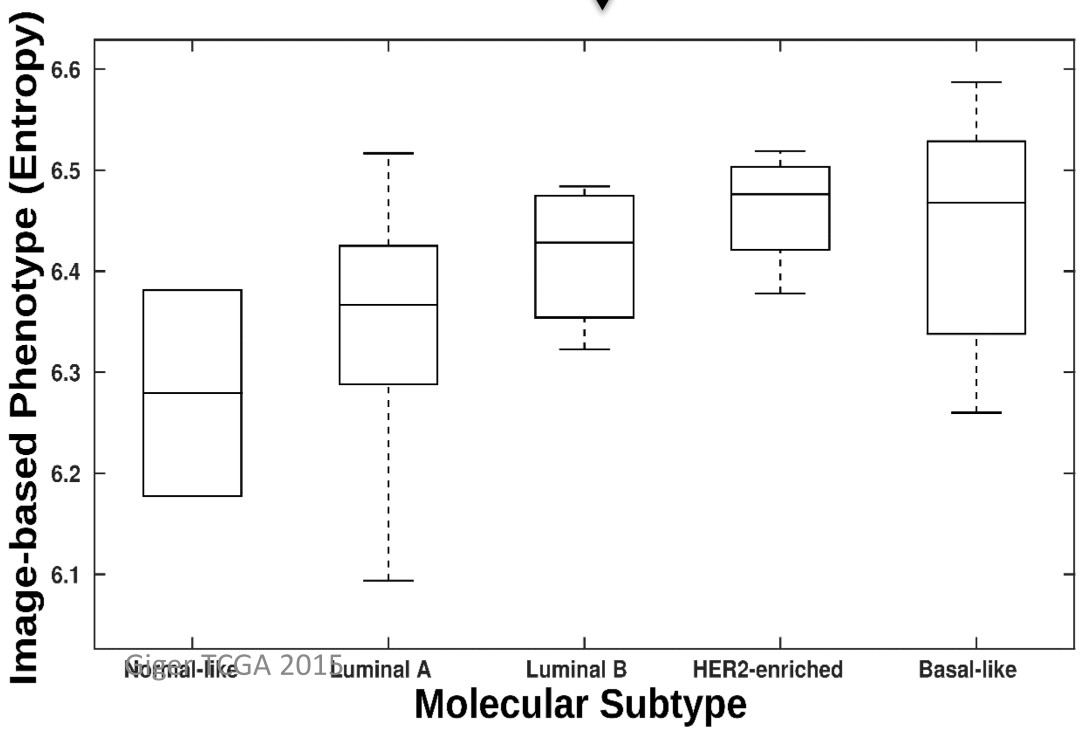


size ≤ 2 cm tumors

Kendall test for trends; p-value=0.0435

2 cm < size ≤ 5 cm

Kendall test for trends; p-value=0.016



Giger TCGA 2015

Molecular Subtyping
from C. Perou

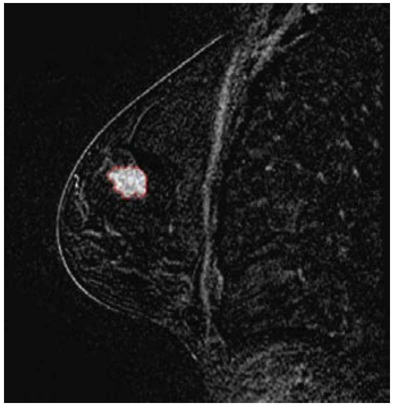
“Virtual biopsy” yielding tumor phenotypes & signatures

Relating Computer-extracted MRI Phenotypes to:

Classification & Association Tasks:

1. Clinical Tumor Status
    1. Tumor Stage
    2. Presence or Absence of Positive Lymph Nodes
2. Molecular Classification & Cancer Subtype
    1. ER- vs. ER+
    2. PR- vs. PR+
    3. Her2- vs. Her2+
    4. Triple Negative vs. Others
3. Risk of Recurrence from multi-gene assays
    1. OncotypeDX
    2. PAM50
    3. MammaPrint
4. Genomic Pathways

Computer analysis of Breast MRIs of tumors



Multi-gene assays of risk of recurrence

Radiomics for "virtual" biopsy

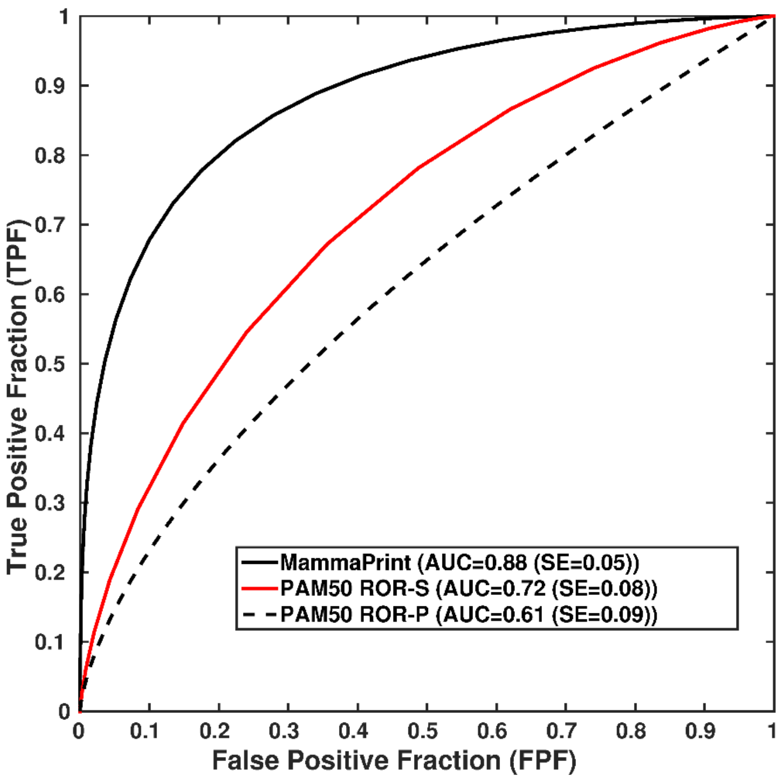
	Good Prognosis Case (left)	**Poor Prognosis Case (right)**
Cancer Subtype	Luminal A	Basal-like
OncotypeDX Range [0, 100]	14.4 (low risk of breast cancer recurrence)	100 (high risk of breast cancer recurrence)
MammaPrint Range [0.848, -0.748]	0.67 (good prognosis)	-0.54 (poor prognosis)
PAM50 ROR-S (Subtype) Range [-7.42, 71.76]	-2.2 (low risk of breast cancer recurrence)	56.3 (high risk of breast cancer recurrence)
PAM50 ROR-P (Subtype+Proliferation) Range [-13.21, 72.38]	0.96 (low risk of breast cancer recurrence)	53.2 (high risk of breast cancer recurrence)
MRI Tumor Size (Effective Diameter) Range [7.8 54.0]	16.8 mm	21.7 mm
MRI Tumor Irregularity Range [0.40 0.84]	0.438	0.592
MRI Tumor Heterogeneity (Entropy) Range [6.00 6.59]	6.27	6.51

Giger TCGA 2015

Radiomics “Virtual Biopsy” & Risk of Recurrence

Research Gene Assay	Correlation coefficient from multiple linear regression analysis		Multiple linear regression model					
	Correlation coefficient	p-value	Phenotypic categories	Feature	Regression coefficient	Standard Error	T Statistics	p-value
MammaPrint	0.55	7.02e-08	—	Constant	-9.5999	3.4665	-2.7693	0.0069798
			Enhancement Texture	Maximum correlation coefficient	3.1407	0.80861	3.884	0.00021048
			Enhancement Texture	Sum average	0.2216	0.084821	2.6125	0.010732
			Size	Effective diameter	-0.019505	0.0060721	-3.2122	0.0018986
OncotypeDX	0.5	1.42e-06	—	Constant	161.52	45.952	3.515	0.00072709
			Kinetic Curve Assessment	Maximum enhanceme nt	10.298	4.0853	2.5208	0.013697
			Enhancement Texture	Maximum correlation coefficient	-200.23	60.938	-3.2858	0.0015108
			Size	Effective diameter	1.8856	0.49819	3.7849	0.00029576
PAM50 (Subtype)	0.56	2.40e-08	—	Constant	491.36	192.17	2.5569	0.012451
			Enhancement Texture	Maximum correlation coefficient	-182.33	44.826	-4.0675	0.00011062
			Enhancement Texture	Sum average	-9.8692	4.7021	-2.0989	0.03898
			Size	Effective diameter	1.2348	0.33661	3.6682	0.00043854

Performance of the MRI Tumor Signatures in the task of predicting Risk of Recurrence (ROC analysis)



Research Gene Assay	Risk of Recurrence Task		Phenotypic Categories
MammaPrint	good prognosis (70)	bad prognosis (14)	Size + Shape + Enhancement variance kinetics
PAM50 (Subtype)	[low+medium] risk of recurrence (69)	[high] risk of recurrence (15)	Size + Enhancement variance kinetics
PAM50 (Subtype + Proliferation)	[low+medium] risk of recurrence (71)	[high] risk of recurrence (13)	Enhancement Texture

ROC curves for leave-one-out LDA classifier using computer-extracted MRI phenotypes as decision variable in the tasks of distinguishing between [low+medium] and high risk levels of recurrence for MammaPrint, PAM50 ROR-S (Subtype), and PAM50 ROR-P (Subtype+Proliferation) from Perou

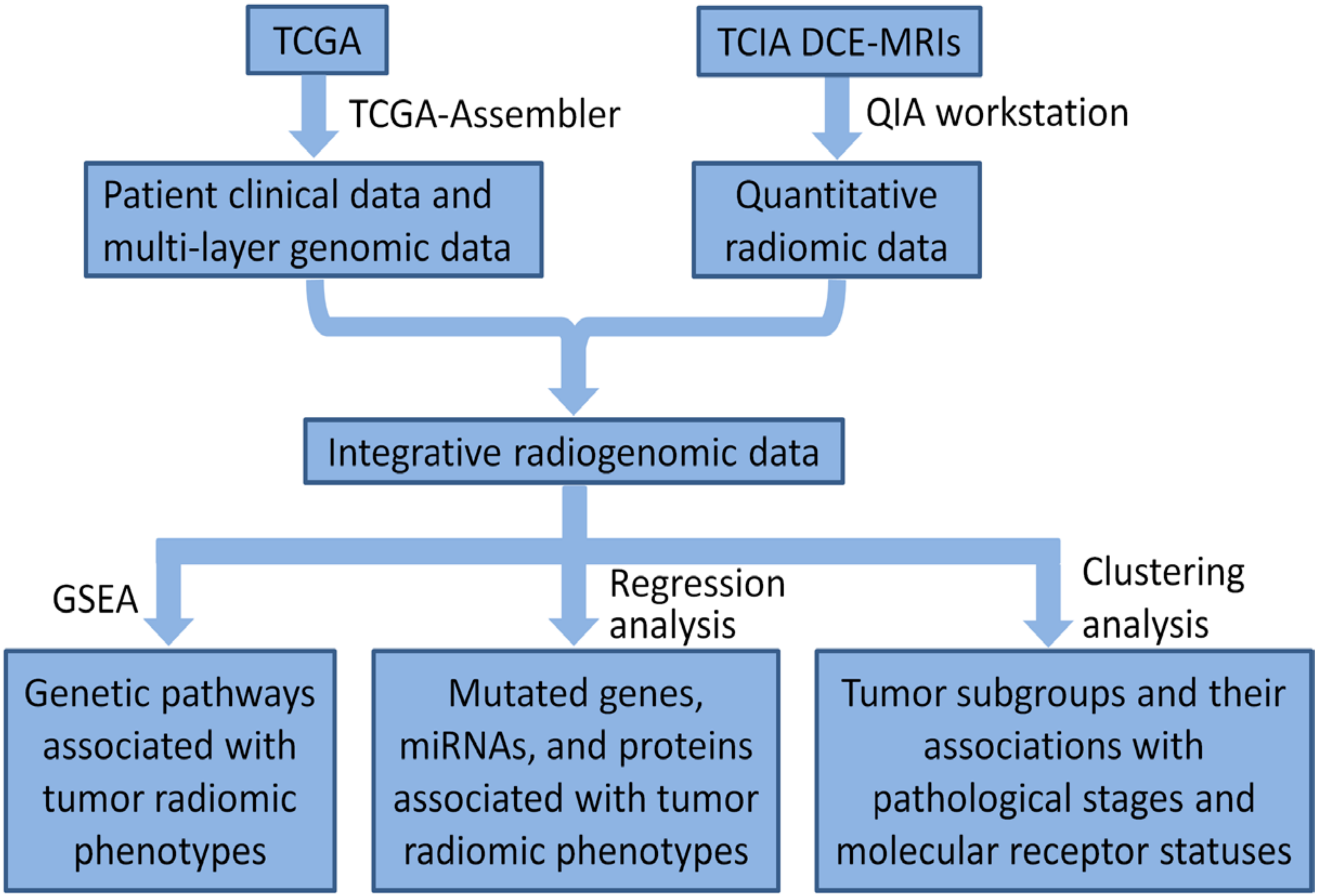
“Virtual biopsy” yielding tumor phenotypes & signatures

Relating Computer-extracted MRI Phenotypes to:

Classification & Association Tasks:

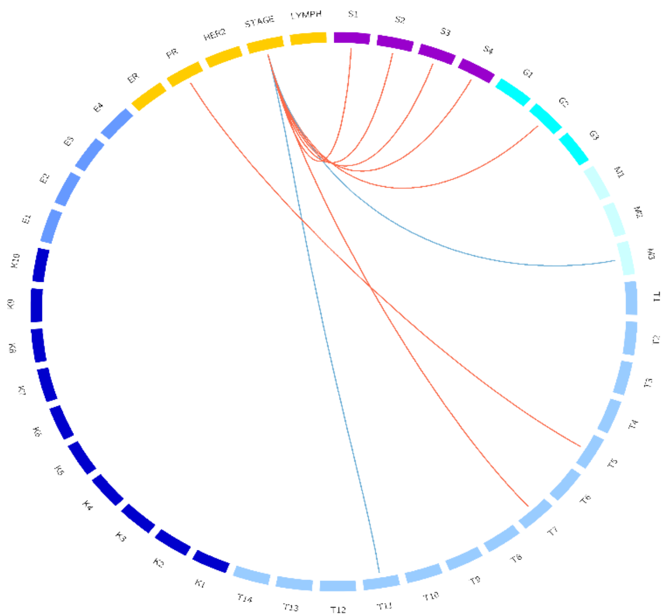
1. Clinical Tumor Status
    1. Tumor Stage
    2. Presence or Absence of Positive Lymph Nodes
2. Molecular Classification & Cancer Subtype
    1. ER- vs. ER+
    2. PR- vs. PR+
    3. Her2- vs. Her2+
    4. Triple Negative vs. Others
3. Risk of Recurrence
    1. OncotypeDX
    2. PAM50
    3. MammaPrint
4. Genomic Pathways

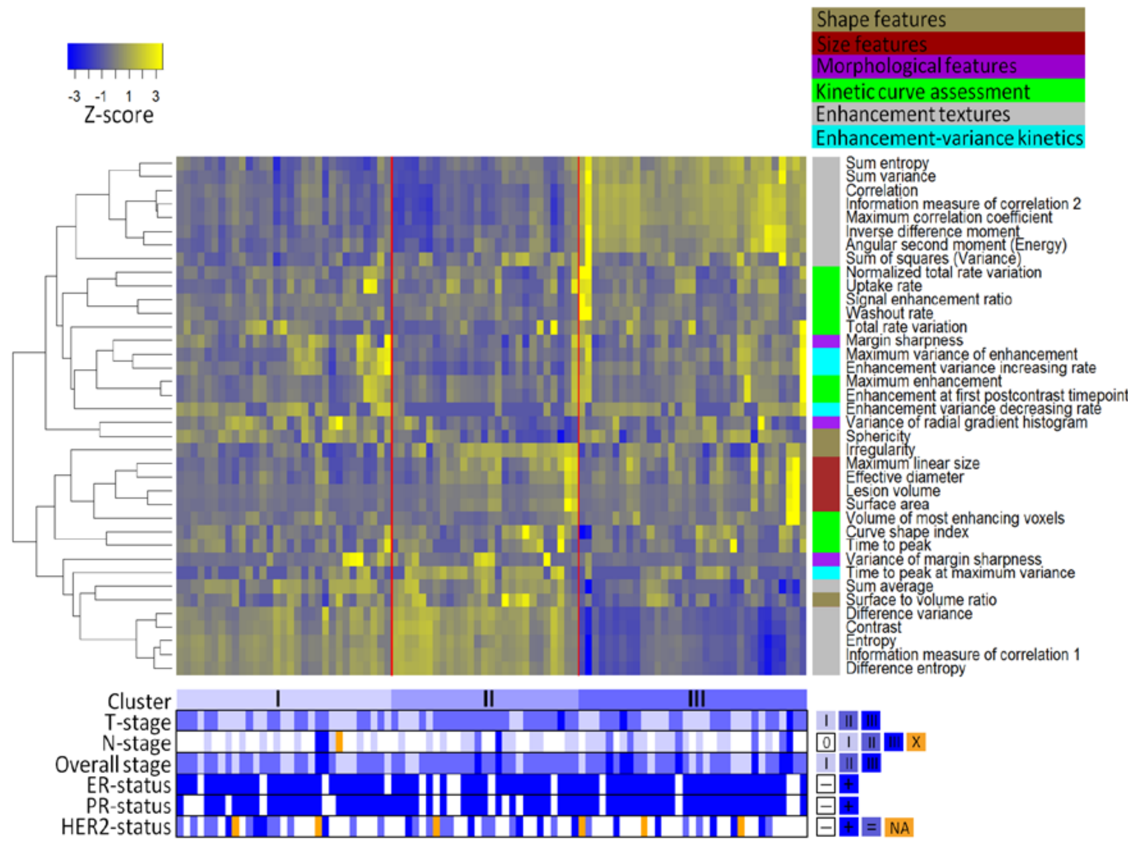
Radiogenomics Flowchart



Giger TCGA 2015

Exploratory Cluster Analysis of the MRI Tumor Phenotypes



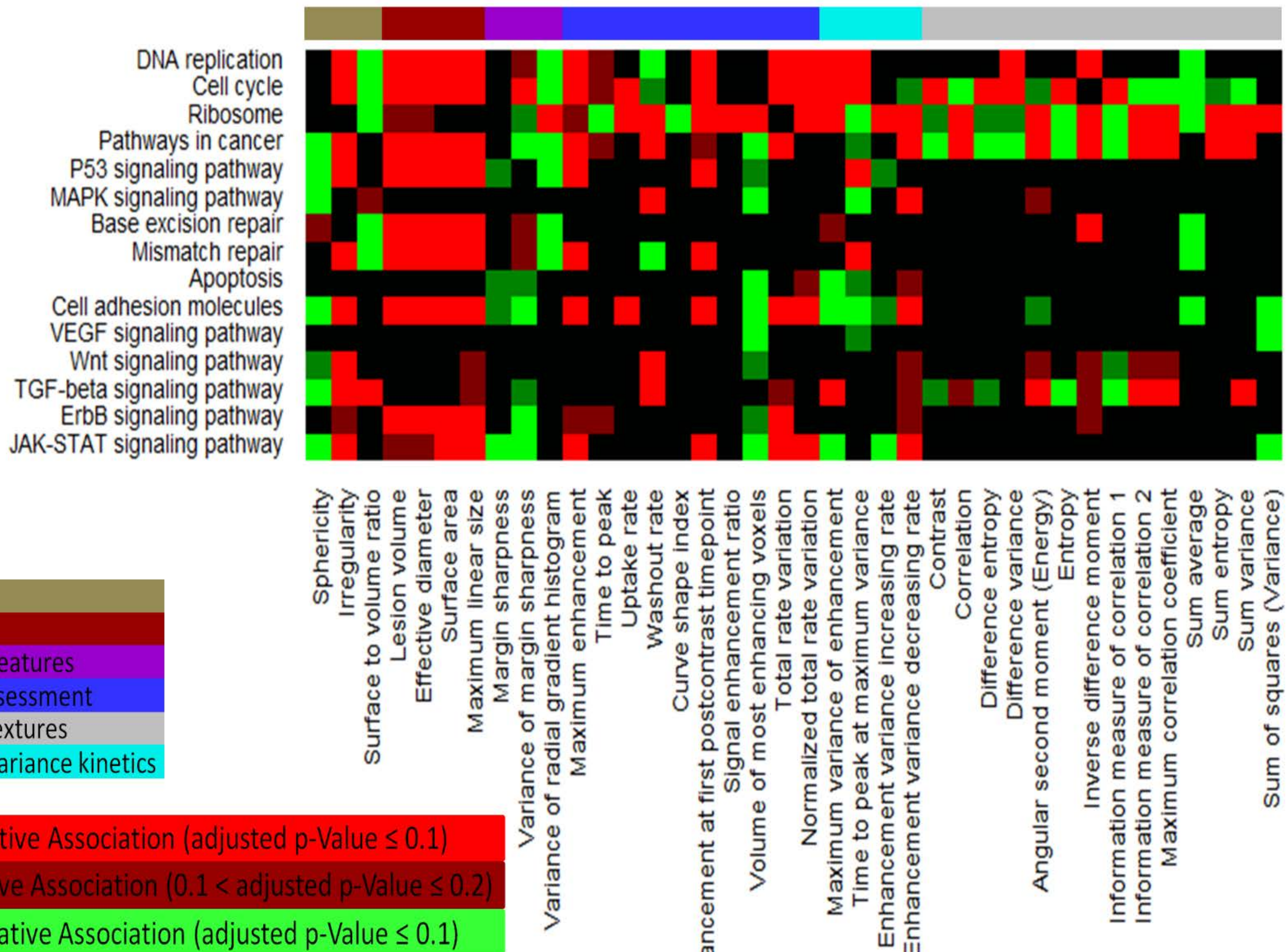


Significant associations between radiomic features and clinical outcomes evaluated by *t*-tests.

Giger TCGA 2015

Radiomics from the MRI tumor "Virtual Biopsy" shows association with Pathway Transcriptional Activities

Ji lab



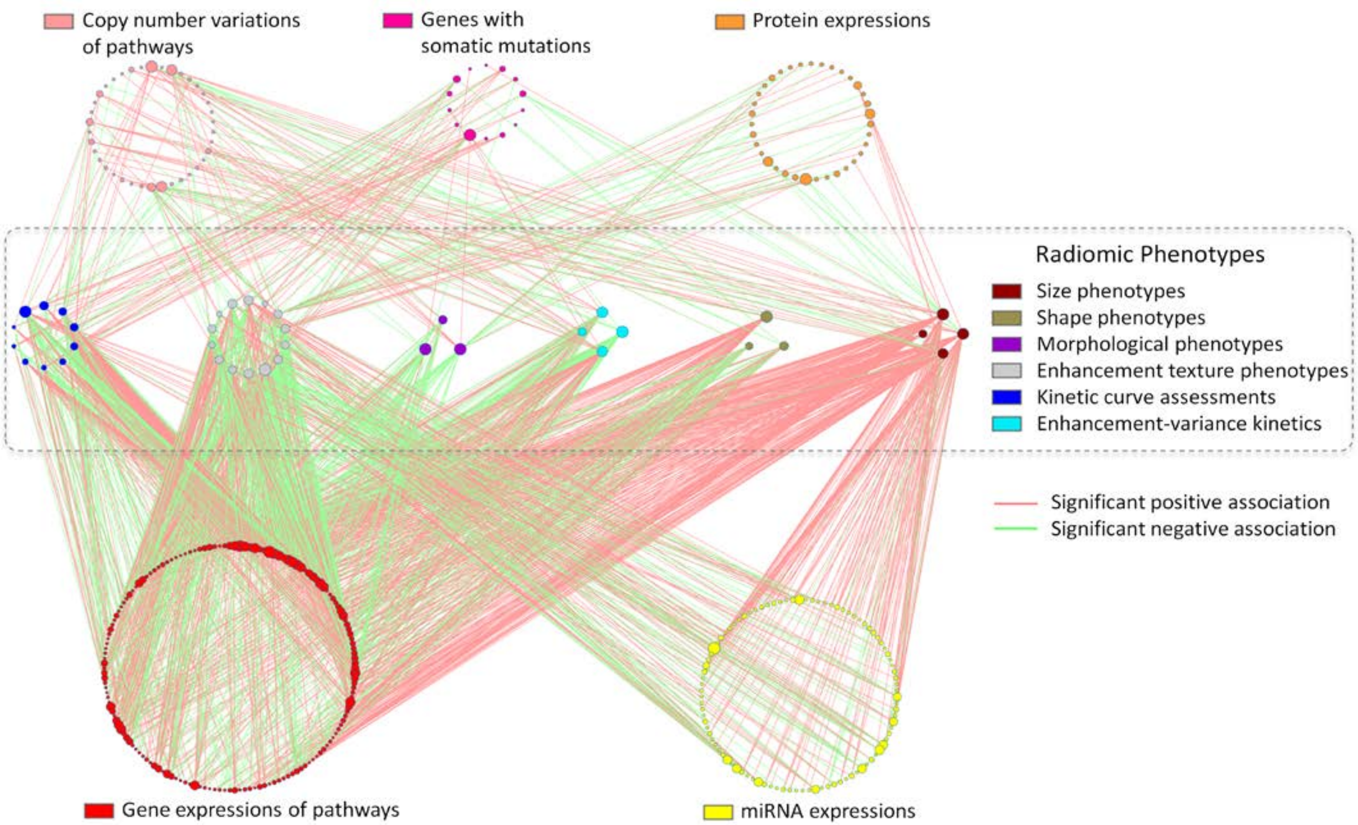
Giger lab

Shape features
Size features
Morphological features
Kinetic curve assessment
Enhancement textures
Enhancement-variance kinetics

Significant Positive Association (adjusted p-Value ≤ 0.1)
Marginal Positive Association (0.1 < adjusted p-Value ≤ 0.2)
Significant Negative Association (adjusted p-Value ≤ 0.1)
Marginal Negative Association (0.1 < adjusted p-Value ≤ 0.2)

Zhu *et al.* submitted

Identified significant associations

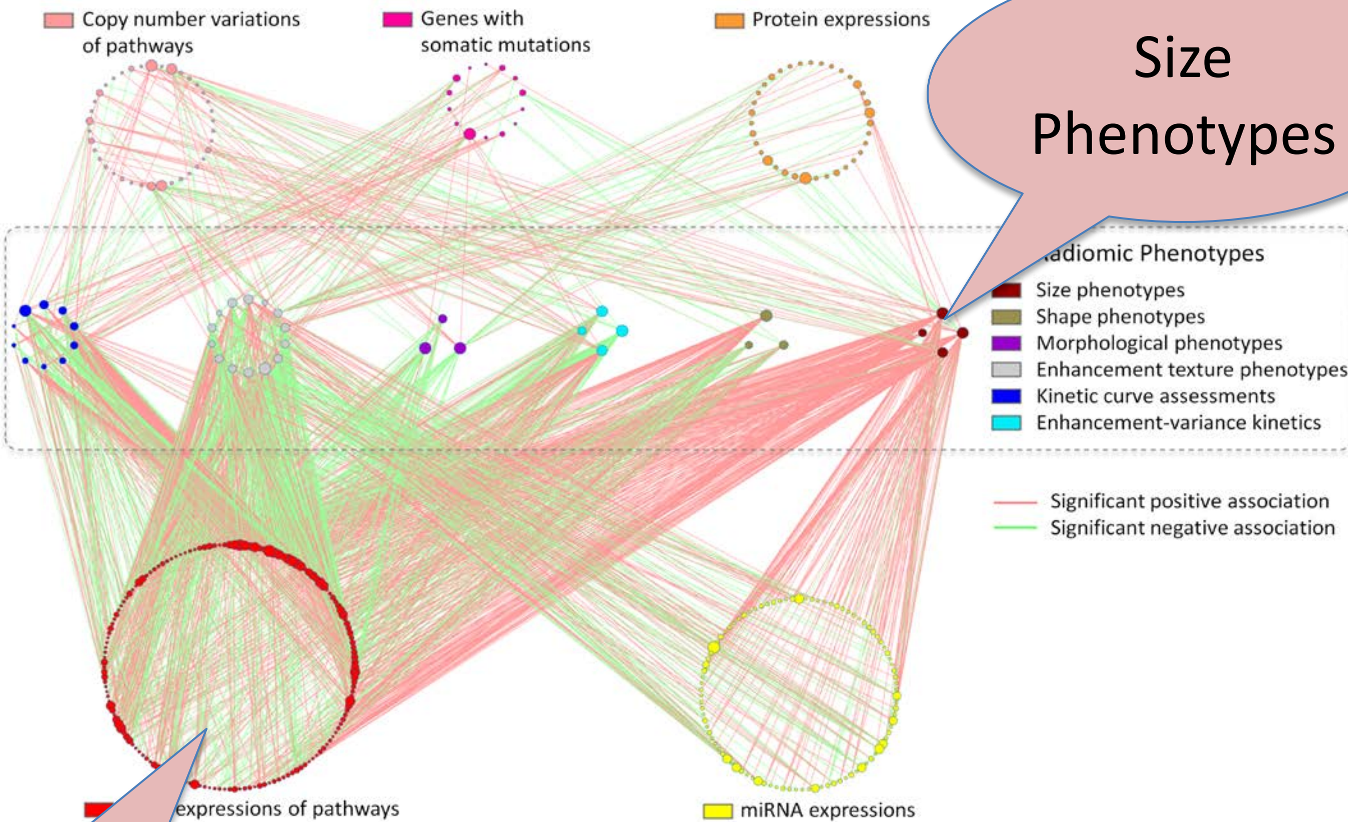


a

	Size phenotypes	Shape phenotypes	Morphological phenotypes	Enhancement texture phenotypes	Kinetic curve assessments	Enhancement-variance kinetics
Gene expressions of pathways	223	150	87	542	334	155
Copy number variations of pathways	29	9	10	25	25	24
Mutated genes	5	3	3	23	35	4
miRNA expressions	126	0	0	142	11	0
Protein expressions	14	0	12	28	1	0

b

Identified significant associations



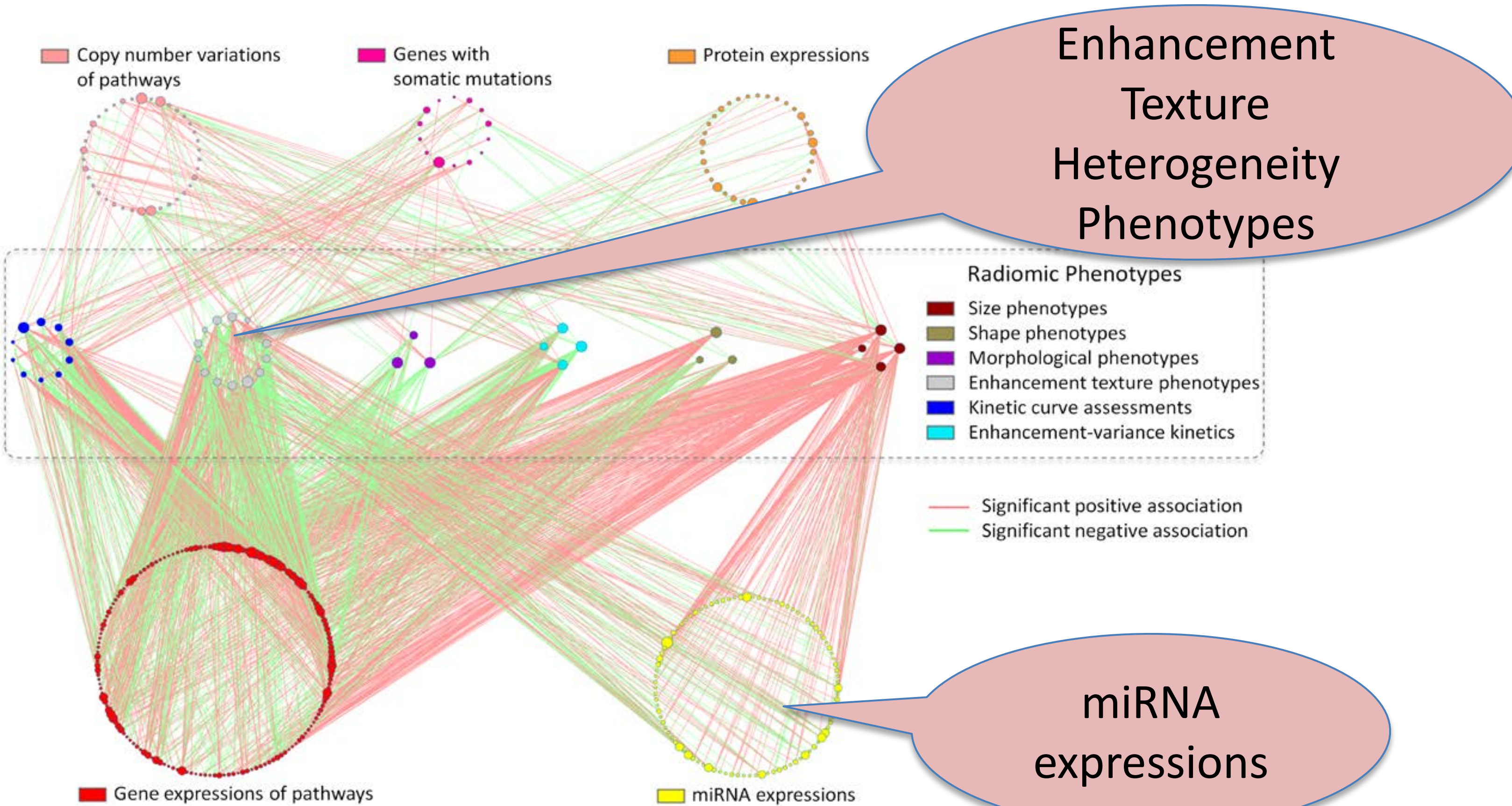
a

Gene expressions of pathways

	Size phenotypes	Shape phenotypes	Morphological phenotypes	Enhancement texture phenotypes	Kinetic curve assessments	Enhancement-variance kinetics
athways	223	150	87	542	334	155
ations of pathways	29	9	10	25	25	24
Mutated genes	5	3	3	23	35	4
miRNA expressions	126	0	0	142	11	0
Protein expressions	14	0	12	28	1	0

b

Identified significant associations



a

	Size phenotypes	Shape phenotypes	Morphological phenotypes	Enhancement texture phenotypes	Kinetic curve assessments	Enhancement-variance kinetics
Gene expressions of pathways	223	150	87	542	334	155
Copy number variations of pathways	29	9	10	25	25	24
Mutated genes	5	3	3	23	35	4
miRNA expressions	126	0	0	142	11	0
Protein expressions	14	0	12	28	1	0

b

Summary & Conclusion

- Computational quantitative MRI analysis shows promise as a means for high-throughput image-based phenotyping and appears to predict breast cancer molecular subtypes

- Radiomics of tumor size and enhancement heterogeneity appear as dominant MRI phenotypes in classifiying tumor subtypes and risk of recurrence.

- Significant associations were identified between the MRI phenotypes (such as tumor size, shape, margin, enhancement texture, blood flow kinetics) and molecular features involved in multiple regulation layers (including DNA mutation, miRNA expression, protein expression, pathway gene expression and copy number variation).

Summary & Conclusion

- Limitations included a small dataset of only 91 cancers
  - TCIA is collecting additional images
  - Investigators are organizing a multi-institutional radiomics network to collect beyond the TCGA/TCIA
- Identification of radiomics of molecular subtypes of breast tumors is expected to allow for virtual biopsies
- Ongoing research involves relating and merging MRI phenotypes with genomic data to develop improved predictive models

Questions

- Is it possible to decide targeted therapy based on imaging-genomics association results?
- Can imaging features inform important genomics features?
- Can integration of imaging and genomics features lead to higher power in prediction?
- Can imaging serve as a virtual biopsy?
  - non-invasive, covers complete tumor, & repeatable

Thank you & please attend our related Workshop & Posters

- Workshop: Imaging Resources for the TCGA: Radiology and Pathology Tools for Enabling Science; May 11; 4-5pm and repeated 5-6pm
- Poster 91
- Poster 79
- Poster 105