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

Giger TCGA 2015
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
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 magnetic 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*
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

Histopathology, Molecular Classification

Genomics

Computer Segmentation of Lesions

Radiologist Descriptors

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

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

Computer Segmentation of Lesions

Radiologist Descriptors

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

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
- Radiologist Descriptors
- Computer Segmentation of Lesions
- 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

Giger TCGA 2015
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

Radiologist Descriptors

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

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

Which correlate and which are complementary???

Lead to Personalized Screening and Personalized Treatment

Giger TCGA 2015
Automated Lesion Segmentation, Feature Extraction [volumetrics, morphological, texture, kinetics] and Estimation of the Probability of Malignancy

Giger et al., RSNA 2010
Dataset

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)

cancergenome.nih.gov
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

MRI 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)

cancergenome.nih.gov
cancerimagingarchive.net
Distribution of the 91 MRI cases

- **ER**:
  - Negative: 14
  - Positive: 77

- **PR**:
  - Negative: 19
  - Positive: 72

- **HER2**:
  - Negative: 72
  - Positive: 19

- **TN**:
  - Negative: 11
  - Positive: 80

Giger TCGA 2015
Distribution of the 91 MRI cases

- Normal-like: 4 cases
- Luminal A: 55 cases
- Luminal B: 10 cases
- HER2-enriched: 5 cases
- Basal-like: 10 cases
Dataset

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

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

Tumor location on MRI determined by consensus of three of the TCIA radiologists
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
Dynamic Contrast-Enhanced MRI & Tumor Segmentation

4D image analysis

Increasing time

Across slices

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

Giger TCGA 2015
Computer-extracted Breast Cancer on MRI (can analyze as a “virtual” biopsy of the tumor)
University of Chicago High-Throughput MRI Phenotyping System

4D DCE MRI images

Radiologist-indicated Tumor Center

Computerized Tumor Segmentation

Computer-Extracted Image Phenotypes (CEIP)

Size
Shape
Morphology
Contrast Enhancement
Texture
Curve
Variance

CAD pipeline = radiomics pipeline

Giger TCGA 2015
University of Chicago High-Throughput MRI Phenotyping System

4D DCE MRI images

- Volume
- Effective diameter
- Maximum linear size
- Surface Area

Computer-Extracted Image Phenotypes

- Contrast Enhancement

Texture
Curve
Variance

CAD pipeline = radiomics pipeline

Giger TCGA 2015
University of Chicago High-Throughput MRI Phenotyping System

4D DCE MRI images

- Sphericity
- Irregularity
- Surface area/volume

CAD pipeline = radiomics pipeline

4D DCE MRI images

Computer-Extracted Image Phenotypes (CEIP)

- Size
- Shape
- Morphology
- Contrast Enhancement
- Texture
- Curve
- Variance
University of Chicago High-Throughput MRI Phenotyping System

4D DCE MRI images

Radiologist-indicated Tumor Center

Computer-Extracted Image Phenotypes (CEIP)

- Size
- Shape
- Morphology
- Contrast Enhancement
- Texture
- Curve
- Variance

CAD pipeline = radiomics pipeline

Margin sharpness
Variance of margin sharpness
Variance of radial gradient histogram

Giger TCGA 2015
University of Chicago High-Throughput MRI Phenotyping System

4D DCE MRI images

Computer-Extracted Image Phenotypes (CEIP)

- Size
- Shape
- Morphology
- Contrast Enhancement
  - Texture
  - Curve
  - Variance

Enhancement heterogeneity & kinetics of the uptake and washout of the contrast agent during the imaging time

CAD pipeline = radiomics pipeline

Giger TCGA 2015
Tumors are Heterogeneous:
Contrast Enhancement Heterogeneity & Kinetics

Heterogeneity of Tumors:

Giger TCGA 2015
4D DCE MRI images

Size

Shape

Morphology

Contrast Enhancement

Texture

Curve

Variance

Contrast enhancement texture characterizing heterogeneity


Giger TCGA 2015
University of Chicago High-Throughput MRI Phenotyping System

4D DCE MRI images

Radiologist-indicated Tumor Center

Computerized Tumor Segmentation


Kinetic curve assessment based on most-enhancing voxels within tumor:
Uptake, washout, curve shape

Size
Shape
Morphology
Contrast Enhancement
Texture
Curve
Variance


Giger TCGA 2015
University of Chicago High-Throughput MRI Phenotyping System For Breast Tumors

4D DCE MRI images

Radiologist-induced Tumor Center

Computerized Tumor Segmentation

Computer-Extracted Image Phenotypes (CEIP)

Size  Shape  Morphology  Contrast Enhancement

Texture  Curve  Variance

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

Giger TCGA 2015
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.
From TCIA Radiomics- Triple Negative Breast Cancers tended to have a more irregular shape, and more heterogeneous in terms of contrast enhancement.
From the TCIA Radiomics -- Enhancement Texture of Tumor Heterogeneity appears Predictive of Molecular Subtype

Molecular Subtyping from C. Perou

Kendall test results for trends; p-value=0.0055
From the TCIA Radiomics -- Enhancement Texture of Tumor Heterogeneity appears **Predictive of Molecular Subtype**

- **size \( \leq 2 \text{ cm} \) tumors**
  - Kendall test for trends; \( p \)-value=0.0435

- **2 cm < size \( \leq 5 \text{ cm} \)**
  - Kendall test for trends; \( p \)-value=0.016

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

<table>
<thead>
<tr>
<th></th>
<th>Good Prognosis Case (left)</th>
<th>Poor Prognosis Case (right)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cancer Subtype</strong></td>
<td>Luminal A</td>
<td>Basal-like</td>
</tr>
<tr>
<td><strong>OncotypeDX</strong></td>
<td>14.4 (low risk of breast cancer recurrence)</td>
<td>100 (high risk of breast cancer recurrence)</td>
</tr>
<tr>
<td><strong>MammaPrint</strong></td>
<td>0.67 (good prognosis)</td>
<td>-0.54 (poor prognosis)</td>
</tr>
<tr>
<td><strong>PAM50 ROR S (Subtype)</strong></td>
<td>-2.2 (low risk of breast cancer recurrence)</td>
<td>56.3 (high risk of breast cancer recurrence)</td>
</tr>
<tr>
<td><strong>PAM50 ROR P (Subtype+Proliferation)</strong></td>
<td>0.96 (low risk of breast cancer recurrence)</td>
<td>53.2 (high risk of breast cancer recurrence)</td>
</tr>
<tr>
<td><strong>MRI Tumor Size</strong></td>
<td>16.8 mm</td>
<td>21.7 mm</td>
</tr>
<tr>
<td><strong>MRI Tumor Irregularity</strong></td>
<td>0.438</td>
<td>0.592</td>
</tr>
<tr>
<td><strong>MRI Tumor Heterogeneity (Entropy)</strong></td>
<td>6.27</td>
<td>6.51</td>
</tr>
</tbody>
</table>

### Multi-gene assays of risk of recurrence

### Radiomics for “virtual” biopsy

Giger TCGA 2015
<table>
<thead>
<tr>
<th>Research Gene Assay</th>
<th>Correlation coefficient from multiple linear regression analysis</th>
<th>Multiple linear regression model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Correlation coefficient</strong></td>
<td><strong>p-value</strong></td>
</tr>
<tr>
<td>MammaPrint</td>
<td>0.55</td>
<td>7.02e-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OncotypeDX</td>
<td>0.5</td>
<td>1.42e-06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAM50 (Subtype)</td>
<td>0.56</td>
<td>2.40e-08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Performance of the MRI Tumor Signatures in the task of predicting Risk of Recurrence (ROC analysis)

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

<table>
<thead>
<tr>
<th>Research Gene Assay</th>
<th>Risk of Recurrence Task</th>
<th>Phenotypic Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>MammaPrint</td>
<td>good prognosis (70)</td>
<td>Size + Shape + Enhancement variance kinetics</td>
</tr>
<tr>
<td>PAM50 (Subtype)</td>
<td>[low+medium] risk of recurrence (69)</td>
<td>Size + Enhancement variance kinetics</td>
</tr>
<tr>
<td>PAM50 (Subtype + Proliferation)</td>
<td>[low+medium] risk of recurrence (71)</td>
<td>Enhancement Texture</td>
</tr>
<tr>
<td></td>
<td>[high] risk of recurrence (15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(13)</td>
<td></td>
</tr>
</tbody>
</table>
“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

Giger TCGA 2015
Radiogenomics Flowchart

TCGA
TCGA-Assembler
Patient clinical data and multi-layer genomic data

TCIA DCE-MRIs
QIA workstation
Quantitative radiomic data

Integrative radiogenomic data

GSEA
Genetic pathways associated with tumor radiomic phenotypes

Regression analysis
Mutated genes, miRNAs, and proteins associated with tumor radiomic phenotypes

Clustering analysis
Tumor subgroups and their associations with pathological stages and molecular receptor statuses

Giger TCGA 2015
Significant associations between radiomic features and clinical outcomes evaluated by $t$-tests.
Radiomics from the MRI tumor “Virtual Biopsy” shows association with Pathway Transcriptional Activities

Zhu et al. submitted

Giger lab
Ji lab
Giger TCGA 2015
Identified significant associations
Identified significant associations

Size Phenotypes
Gene expressions of pathways

<table>
<thead>
<tr>
<th>Pathways</th>
<th>Size phenotypes</th>
<th>Shape phenotypes</th>
<th>Morphological phenotypes</th>
<th>Enhancement texture phenotypes</th>
<th>Kinetic curve assessments</th>
<th>Enhancement-variance kinetics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy number variations</td>
<td>223</td>
<td>150</td>
<td>87</td>
<td>542</td>
<td>334</td>
<td>155</td>
</tr>
<tr>
<td>of pathways</td>
<td>29</td>
<td>9</td>
<td>10</td>
<td>25</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>Mutated genes</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>23</td>
<td>35</td>
<td>4</td>
</tr>
<tr>
<td>miRNA expressions</td>
<td>126</td>
<td>0</td>
<td>0</td>
<td>142</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Protein expressions</td>
<td>14</td>
<td>0</td>
<td>12</td>
<td>28</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Identified significant associations

Enhancement
Texture
Heterogeneity
Phenotypes

Radiomic Phenotypes
- Size phenotypes
- Shape phenotypes
- Morphological phenotypes
- Enhancement texture phenotypes
- Kinetic curve assessments
- Enhancement-variance kinetics

Significant positive association
Significant negative association

Gene expressions of pathways
miRNA expressions

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