The “-Omic” World I

- **Definition:**
  - A field of study in **biology** ending in –omics (genomics, transcriptomics, proteomics or metabolomics)

- **Objective:**
  - Collective **characterization** and **quantification** of pools of biological molecules that translate into the structure, function, and dynamics of an organism(s)

The “-Omic” World II

**Omics-Based Test Development Process**

- According to the Institute of Medicine (IOM):
Radiogenomic Modeling of Rectal Bleeding in Prostate cancer

Coates et al., RO, 2015

Radiogenomics assessment via PCA visualization

Identification of robust biomarkers for RP using Proteomics with limited samples (n= 3x3)

Oh, Craft et al., JPR, 2011

Radioproteomics in Lung Cancer

Lee et al., Med Phys, 2015
Radiation Response as Pan-Omics

Clinical Factors
- Demographics
- Histology
- Dose Prescription
- Treatment Technique
- Planning Data
- Functional imaging
- DoSg-Volume Metrics

Physical Factors
- DNA damage detection and repair genes
- Tumor Stage
- Tumor Volume
- Fibrotic and Inflammatory Cytokines
- Anti-oxidant Enzymes

Panomics
Integration of physics (radiomics) and biology (genomics)

Lung Cancer Jamboree

Imagin

Biomarkers

Fibrosis
Pneumonit

Radiomics

- A ‘new’ form of –omics
- Quantitative information from multi-imaging modalities (PET, CT, MRI, etc) could be related to biological and clinical endpoints (Lambin et al, 2012)
- In oncology, it is decoding the Tumor Phenotype with Non-Invasive Imaging (radiomics.org)
An image worth thousand(s) words

Our early radiomics work

More recently: HNC PET

Common radiomics features
PET/CT from NSCLC local tumor

SUVmax = 13.7

IVH

Texture map

Radiomics PET/CT model

Textures VS CTCv4 fibrosis score

Radiologists' assessments
Textures VS dose/time

Textures VS biomarkers

PET/MR fusion for Sarcoma mets to the lungs

Methodology
Fused versus Separate Scans

**QUESTION**
Do texture features extracted from **FUSED scans** provide better assessment of tumor aggressiveness than those extracted from **SEPARATE scans**?

**FUSION EXAMPLE**

**MULTIVARIABLE ANALYSIS**
(Identifying optimal parsimonious model)

- **FDG-PET**
- **FDG-PET/T2FS**
- **MRI T2FS**

**MULTIVARIABLE MODELING**
(Logistic regression [LR])

- **Set balanced model**
  - Prediction from balanced training set
- **Varying initialization**
- **Maximization of AUC**
- **ROCAUC analysis**
  - Correction for small sample size effect
  - Visible of best model

**FINAL MODEL COMPUTATION**
(LR coefficients)

**TESTING DATA SIMULATION**
(Imbalance-adjusted BOOTSTRAPPING)

**FEATURE SET REDUCTION**

- **INITIAL FEATURE SET**
  - 41 textures * 240 extraction parameter combinations:
  - ~10,000 texture-parameter features

Goal: Allow the creation of a feature set balance between predictive power and maximal information (Gain)

- **Part 1**
  - Prognostic value

\[ Ge_{j} = y \cdot \rho_{j} \cdot [E(x_{j},y)] + \delta_{j} \]

- **Part 2**
  - Interdependence of feature with already chosen features

\[ \sum_{j=1}^{k} \frac{(2f-k+1)}{f(f+1)} \cdot \rho(C(x_{j},y)) \]

- **Part 3**
  - Interdependence of feature with features not yet removed

\[ \sum_{j=1}^{k} \frac{(2f-k+1)}{f(f+1)} \cdot \rho(C(x_{j},y)) \]

**FEATURE SET REDUCTION**
Part 1 – Prognostic value

\[ r_{i}(x,y) : \text{Spearman’s rank correlation between feature \( x \) and outcome \( y \) (Lung Mets)} \]

For each bootstrap sample, calculate a new \( r_{i}(x,y) \) from the training set. Repeat for 1000 bootstrap samples and record the mean.
FEATURE SET REDUCTION

Part 2 – Interdependence with selected features

\[ \text{PIC} = 1 - \text{MIC} \]

Potential Information coefficient

\[ \text{MIC} : \frac{\sum_{i=1}^{n} P(x_i, y_i) - \left( \frac{\sum_{i=1}^{n} P(x_i) \cdot P(y_i)}{n} \right)}{ \sqrt{ \left( \frac{\sum_{i=1}^{n} (P(x_i) - \frac{1}{n})^2}{n-1} \right) \cdot \left( \frac{\sum_{i=1}^{n} (P(y_i) - \frac{1}{n})^2}{n-1} \right) } } \]

MIC: Maximal Information coefficient. Has the ability to capture any simple or complex relationship types of association between two variables (independent to the modeled outcome).

---

FEATURE SET REDUCTION

Part 3 – Interdependence with unselected features

\[ \text{PIC} = 1 - \text{MIC} \]

Potential Information coefficient

\[ \text{MIC} : \frac{\sum_{i=1}^{n} P(x_i, y_i) - \left( \frac{\sum_{i=1}^{n} P(x_i) \cdot P(y_i)}{n} \right)}{ \sqrt{ \left( \frac{\sum_{i=1}^{n} (P(x_i) - \frac{1}{n})^2}{n-1} \right) \cdot \left( \frac{\sum_{i=1}^{n} (P(y_i) - \frac{1}{n})^2}{n-1} \right) } } \]

MIC: Maximal Information coefficient. Has the ability to capture any simple or complex relationship types of association between two variables (independent to the modeled outcome).

---

Model Order Selection

Most parsimonious model (set of features)

COMPLETE PREDICTION MODEL

Identified set of features:

1. PET/T2FS -- SZE: Small Zone Emphasis texture extracted on fused PET/T2FS scans
2. PET/T1 -- ZSV: Zone Size Variance texture extracted on fused PET/T1 scans
3. PET/T1 -- HGZE: High Gray-Level Zone Emphasis extracted on PET/T1 scans
4. PET/T2FS -- HGRE: High Gray-Level Run Emphasis extracted on PET/T2FS scans

Final multivariable model response:

\[ g(x) = 255 \cdot \text{PET/T2FS} + 5360 \cdot \text{PET/T1} + 1.75 \cdot \text{PET/T1} - 1.60 \cdot \text{PET/T2FS} + 5.50 \]

\[ \epsilon(x) \cdot r(y) - (1 - \epsilon(x)) \cdot \frac{\mu(x)}{\epsilon(x)} \]
Conclusions

• Treatment outcomes are multifactorial (Pan-Omics)
  – Combination of physical (radiomics) and biological (radiogenomics) factors
• Radiomics is an essential element of the Pan-Omics world and constitute a powerful tool to interrogate wealthy imaging information
  – Single and multiple modalities
    • Separate and fused
• Radiomics involves two main steps
  – Robust feature extraction
  – Robust modeling

ACKNOWLEDGMENTS

• Joseph Desey, PhD
• Carolyn R. Freeman, MBBS
• Jan Seuntjens, PhD
• Sonia R. Shamene, MD
• Sangkyu Lee, PhD candidate
• Martin Carrier-Vallieres, PhD Candidate