

**AAPM 2016**   **AAPM 2016**   **AAPM 2016**

**CHALLENGES**  
IN  
**RADIOMICS AND BIG DATA**

 THE UNIVERSITY OF CHICAGO

Karen Drukker

---

---

---

---

---

---

---

---

 THE UNIVERSITY OF CHICAGO

**ACKNOWLEDGEMENTS**

- Lorenzo Pesce
- Maryellen Giger

**FUNDING**

- Supported over the years by The University of Chicago Dean Bridge Fund, grants from the NIH (NCI, NIBIB, NIAMS), the DOD, the DOE, and the Segal Family Foundation, the UC Comprehensive Cancer Center, the SPORC, and the ITM/CTSA

**CONFLICT OF INTEREST STATEMENT OF PRESENTER**

- Karen Drukker receives royalties from Hologic



---

---

---

---

---

---


---

---

**PURPOSE**

- To summarize some of the statistical challenges in radiomics, genomics, radiogenomics, and big data from the perspective of a novice

<sup>†</sup>In this talk radiogenomics refers to the combination of radiomic features and genomic data, different from the use of this term in radiation oncology



---

---

---

---

---

---

---

---

**INTRODUCTION**  
**WHAT IS RADIOMICS?**

Radiology

ORIGINAL RESEARCH • SPECIAL REPORT

### Radiomics: Images Are More than Pictures, They Are Data<sup>1</sup>

Robert J. Gillies, PhD  
Paul E. Kinahan, PhD  
Heung Hock, MD, PhD, Dr(h)

In the past decade, the field of medical image analysis has grown exponentially, with an increased number of pattern recognition tools and an increase in data set sizes. These advances have facilitated the development of processes for

R.J. Gillies et al, Radiology, 278, p563-477, 2016

---

---

---

---

---

---

---

---

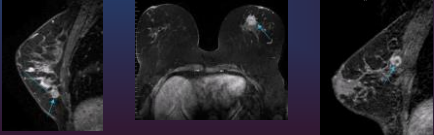
---

---

**INTRODUCTION**  
**WHAT IS RADIOMICS?**

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

Image Data



---

---

---

---

---

---

---

---

---

---

**INTRODUCTION**  
**DEFINITIONS**

- Radiomics:** High throughput conversion of images to mineable data

↓

This data can be viewed as descriptors (i.e., phenotypes) of tumors and "normal" tissue

- Big Data:** The exponential growth in the numbers of patients and the data elements being harvested from each is known colloquially as "big data"

R.J. Gillies et al, Radiology, 278, p563-477, 2016

---

---

---

---

---

---

---

---

---

---

## INTRODUCTION

### DEFINITIONS

#### Phenotype:

- The observable physical or biochemical characteristics of an organism, as determined by both genetic makeup and environmental influences
- The expression of a specific trait, such as stature or blood type, based on genetic and environmental influences



---

---

---

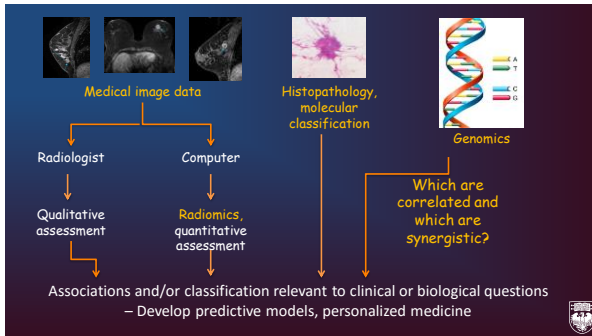
---

---

---

---

---



---

---

---

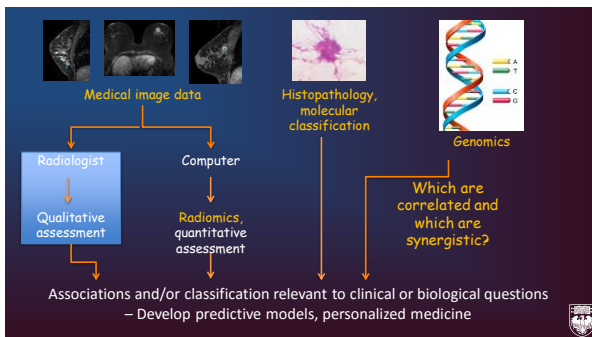
---

---

---

---

---



---

---

---

---

---

---

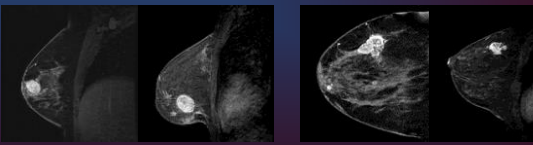
---

---


**INTRODUCTION**  
**EXAMPLES OF IMAGE-BASED PHENOTYPES**

Radiologist-assessed **qualitative** semantic phenotypes

**Shape**



Round-oval                      Irregular



---

---

---

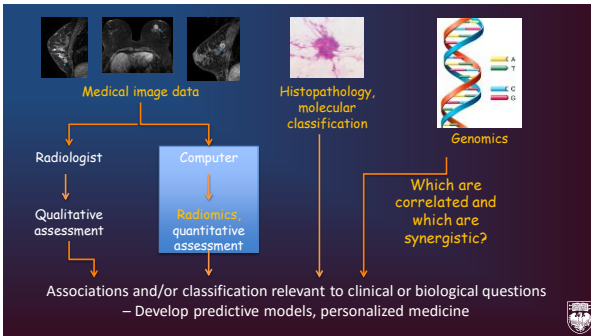
---

---

---

---

---



---

---

---

---

---

---

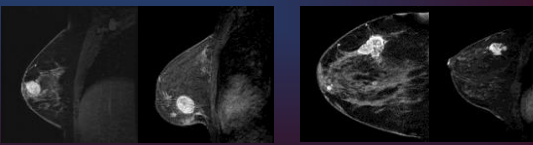
---

---


**INTRODUCTION**  
**EXAMPLES OF IMAGE-BASED PHENOTYPES**

In radiomics, we obtain **quantitative** computer-extracted image-based phenotypes (sometimes also referred to as agnostic phenotypes)

**Shape**



Sphericity: 0.80; 0.85                      Irregularity: 0.65; 0.78



---

---

---

---

---

---

---

---

**INTRODUCTION**

**EXAMPLES OF IMAGE-BASED PHENOTYPES**

- Characterization of tumor and/or parenchyma texture
- Margin irregularity and sharpness
- CT: Lesion size (RECIST, volume)
- MRI: Kinetic characterization (uptake, washout)
- Nuclear medicine: SUV

↓

"What's in a name -  
A rose by any other name would smell as sweet"  
(Shakespeare)

Quantitative Imaging  
(QIN, QIBA)

Computer-Aided  
Diagnosis

Radiomics

---

---

---

---

---

---

---

---

---

---

**BUT NOW ON TO THE ACTUAL TALK**

**PURPOSE**

- To summarize some of the **statistical** challenges in radiomics, genomics, radiogenomics, and big data from the perspective of a novice

<sup>†</sup>In this talk radiogenomics refers to the combination of radiomic features and genomic data, different from the use of this term in radiation oncology

---

---

---

---

---

---

---

---

---

---

**INTRODUCTION**

**Two types of research**

- Studies that can get by with small sample sizes

Have luxury of being able to probe same system in multiple *independent* ways, e.g., in molecular biology

- Studies that use large data sets and rely mostly or wholly on statistics

This includes us! Radiomics, genomics, radiogenomics

---

---

---

---

---

---


---

---

---

---

**INTRODUCTION**  
**CENTRAL QUESTION**



Know when your numbers are significant

"..... must grasp basic statistics ... or sloppy science will continue to grow"

DL Vaux, Nature, 492, p180-81, 2012

---

---

---

---


---

---

---

---

**INTRODUCTION**  
**CENTRAL QUESTION**



Know when your numbers are significant

- Most scientist have been taught some statistics along the way
  - But often the type of statistics is not relevant to their current work
  - And once in the lab, people generally just do what everyone else does, without necessarily understanding why

DL Vaux, Nature, 492, p180-81, 2012

---

---

---

---

---

---

---

---

**INTRODUCTION**  
**CHALLENGES IN RADIOMICS**

- Reproducibility
- "Big Data"
- Data sharing
- Standardization/harmonization

RJ Gillies et al, Radiology, 278, p563-477, 2016

---

---

---

---

---

---


---

---

**INTRODUCTION**  
**CHALLENGES IN RADIOMICS**

- Reproducibility
- "Big Data"
- Data sharing
- Standardization/harmonization

RJ Gillies et al, Radiology, 278, p563-477, 2016



---

---

---

---

---

---

---

---


**INTRODUCTION**  
**REPRODUCIBILITY**

What is the extent of the problem?

A 2012 study of 53 landmark papers in basic cancer research was able to replicate the original results of just 6 of these studies

GC Begley et al, Nature, 483: 531-33, 2012

RJ Gillies et al, Radiology, 278, p563-477, 2016



---

---

---

---

---

---

---

---


**INTRODUCTION**  
**REPRODUCIBILITY**

What is the extent of the problem?

A 2012 study of 53 landmark papers in basic cancer research was able to replicate the original results of just 6 of these studies

GC Begley et al, Nature, 483: 531-33, 2012

RJ Gillies et al, Radiology, 278, p563-477, 2016



---

---

---

---

---

---

---

---

**INTRODUCTION**

**BIG DATA**

What are the challenges?

- Large data sets
  - Not from carefully controlled experiments
  - Non-uniformity in equipment manufacturer, imaging protocol, imaged population etc



---

---

---

---

---

---

---

---

**OUTLINE OF THE REST OF THIS PRESENTATION**

- Hypothesis testing
- More on correlation and causality
- Reproducibility ↔ *“Big Data”*
  - Radiomics
  - Genomics
  - Radiogenomics
- Discussion



---

---

---

---

---

---

---

---

**HYPOTHESIS TESTING**



---

---

---

---

---

---

---

---



### HYPOTHESIS TESTING

- Hypothesis testing is like a criminal trial: A defendant is considered **not guilty** until proven otherwise

Null-hypothesis  $H_0$

	$H_0$ is true Truly not guilty	$H_1$ is true Truly guilty
Accept null hypothesis Acquittal	Right decision	Wrong decision Type II Error
Reject null hypothesis Conviction	Wrong decision Type I Error	Right decision




---

---

---

---

---

---

---

---

### HYPOTHESIS TESTING

Never confuse Type I and II errors again:

Just remember that the Boy Who Cried Wolf caused both Type I & II errors, in that order.

First everyone believed there was a wolf, when there wasn't. Next they believed there was no wolf, when there was.

Substitute "effect" for "wolf" and you're done.

Kudos to @danoliver for the thought. Illustration by Francis Barlow "De pastoris puero et agricolis" (1687), Public Domain. Via wikimedia.org




---

---

---

---

---

---

---

---

### HYPOTHESIS TESTING

- How do we decide whether to accept or reject the null-hypothesis?

P-values are commonly used *With  $p < 0.05$  denoting a statistically significant difference (for a single comparison)*

Q: Why do so many colleges and grad schools teach  $p=0.05$ ?

A: Because that's still what the scientific community and journal editors use

Q: Why do so many people still use  $p=0.05$ ?

A: Because that's what they were taught in college or grad school

The American Statistician, 70, 129-133, 2016




---

---

---

---

---

---

---

---

**HYPOTHESIS TESTING**

**ASA Statement on Statistical Significance and P-Values**

Absence of evidence is not evidence of absence  
Altman DG and Bland JM, British medical journal, 311, 485, 1995

The statistical crisis in science  
Geldman A and Loken E, American Scientist, 2014

The fallacy of the null-hypothesis significance test  
Kozaboom WM, Psychological Bulletin, 57, 416-428, 1960

False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as statistically significant  
SimmonsJP et al, Psychological Science, 22, 1399-1366, 2011

Why most published research findings are false  
Ioannidis JP, PlosOne, 2, e124, 2005

The American Statistician, 70, 129-133, 2016



---

---

---

---

---

---

---

---

---

---

**HYPOTHESIS TESTING**

**ASA Statement on Statistical Significance and P-Values**

1. p-values can indicate how incompatible the data are with a specified statistical model
2. A p-value does not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone
3. Scientific conclusions and business decisions should not be based on whether a p-value passes a specific threshold
4. Proper inference requires full reporting and transparency
5. A p-value, or statistical significance, does not measure the size of an effect or the importance of a result
6. By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis



---

---

---

---

---

---

---

---

---

---

**HYPOTHESIS TESTING**

**ASA Statement on Statistical Significance and P-Values**

1. p-values can indicate how incompatible the data are with a specified statistical model
2. A p-value does not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone
3. Scientific conclusions and business decisions should not be based on whether a p-value passes a specific threshold
4. **Proper inference requires full reporting and transparency**
5. A p-value, or statistical significance, does not measure the size of an effect or the importance of a result
6. By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis



---

---

---

---

---

---

---

---

---

---

### HYPOTHESIS TESTING

#### MULTIPLE HYPOTHESIS TESTING

- Often a variety of tests for a variety of possible effects are applied to a single data set and only those yielding a significant result are reported

**WRONG**

Cherry-picking promising findings

Selective inference  
 Data dredging  
 Significance questing  
 Significance chasing  
 P-hacking

One needs multiplicity correction procedures that control the family wise error rate or the false discovery rate




---

---

---

---

---

---

---

---

### HYPOTHESIS TESTING

#### MULTIPLE HYPOTHESIS TESTING

- Multiplicity correction procedures
  - 'Traditional' methods to control the family-wise error rate such as Bonferroni-Holm seen by many as too conservative
    - The test fails to reject the null-hypothesis (that there is no difference) when there really is a difference. (type II error)
  - Methods to control the false discovery rate (e.g., the Benjamini-Hochberg procedure) may be more useful, but perhaps too lax?
    - The null-hypothesis is rejected more frequently, but at the cost of indicating a statistically significant difference when there is none. (type I error)




---

---

---

---

---

---

---

---

### MORE ON CORRELATION AND CAUSALITY




---

---

---

---

---

---

---

---

**MORE ON CORRELATION AND CAUSALITY**

Association

POINTS OF SIGNIFICANCE

**Association, correlation and causation**

Correlation Causation

N Altman, M Krzywinski, Nature Methods, 12, p899-900, 2015

---

---

---

---

---

---

---

---

**MORE ON CORRELATION AND CAUSALITY**

Association = Dependence

Associated Not associated

Correlation Causation

N Altman, M Krzywinski, Nature Methods, 12, p899-900, 2015

---

---

---

---

---

---

---

---

**MORE ON CORRELATION AND CAUSALITY**

Association

Trends (linear or increasing/decreasing) =

Uncorrelated Correlated

Correlation Causation

N Altman, M Krzywinski, Nature Methods, 12, p899-900, 2015

---

---

---

---

---

---

---

---

### MORE ON CORRELATION AND CAUSALITY

Association

Null hypothesis: 2 methods/variables are not linearly related

A high correlation does not necessarily mean agreement

Correlation Causation

JM Bland, DG Altman, Lancet i: p307-10, 1986

---

---

---

---

---

---

---

---

---

---

### MORE ON CORRELATION AND CAUSALITY

Association

But even if the points lie along the equality line

Is this agreement "good enough"?

Correlation Causation

JM Bland, DG Altman, Lancet i: p307-10, 1986

---

---

---

---

---

---

---

---

---

---

### MORE ON CORRELATION AND CAUSALITY

Association

More insightful way of assessing agreement

Whether this is "good enough" depends on clinical task bias

Expected range in differences between measurements

Correlation Causation

JM Bland, DG Altman, Lancet i: p307-10, 1986

---

---

---

---

---

---

---

---

---

---

### MORE ON CORRELATION AND CAUSALITY

Association

Very different datasets may have similar/identical correlation coefficients

Correlation also depends on

- Noise
- Sample size

Correlation

Causation

N Altman, M Krzywinski, Nature Methods, 12, p899-900, 2015

---

---

---

---

---

---

---

---

---

---

### MORE ON CORRELATION AND CAUSALITY

Association

Spurious correlation coefficients in random (independent) data

Distribution of correlation coefficients

95% confidence intervals of correlation coefficients

Correlation

Causation

N Altman, M Krzywinski, Nature Methods, 12, p899-900, 2015

---

---

---

---

---

---

---

---

---

---

### MORE ON CORRELATION AND CAUSALITY

Association

Why is this important to us?

- When the number of features is large with respect to the sample size, large but spurious correlations frequently occur
- When there is a large number of observations, small and substantively unimportant correlations may appear statistically significant

*"Small data" pilot studies*

*"Big data": standard of care*

Correlation

Causation

N Altman, M Krzywinski, Nature Methods, 12, p899-900, 2015

---

---

---

---

---

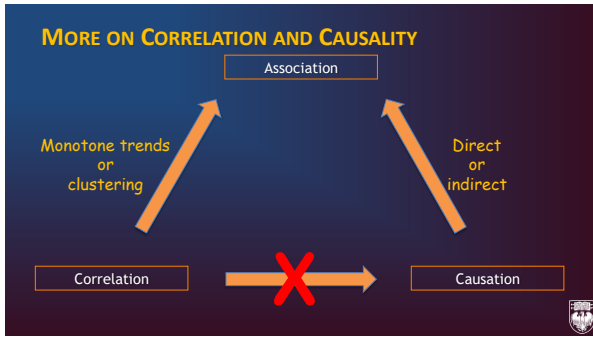
---

---

---

---

---



---

---

---

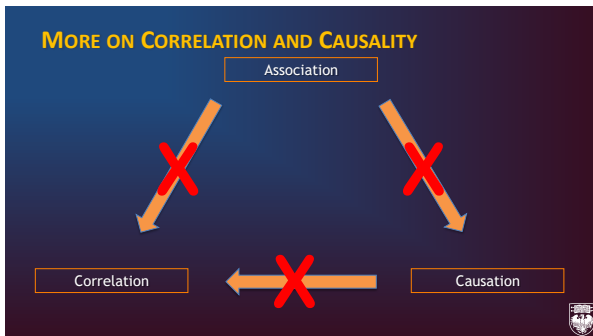
---

---

---

---

---



---

---

---

---

---

---

---

---



---

---

---

---

---

---

---

---





### RADIOMICS

#### 'SMALL DATA' CHALLENGE: OVERFITTING

Many quantitative computer-extracted image-based features usually describe a single physical characteristic

- The 'machine' is able to extract much more, and perhaps more useful, information than the 'human' (hundreds or even thousands of quantitative image-based features, i.e., phenotypes)
- In practice, often data sets of limited size are available for research and one needs to be able to obtain a realistic estimate of performance in the 'real world'




---

---

---

---

---

---

---

---

### RADIOMICS

#### 'SMALL DATA' CHALLENGE: OVERFITTING

Many quantitative computer-extracted image-based features usually describe a single physical characteristic

- Ideally use large training, calibration, and independent test data sets
  - In practice, modest-sized data sets using cross-validation or bootstrapping can give reasonable performance estimates when used properly
- Reduce the number of computer-extracted features through supervised feature selection or unsupervised dimension reduction/clustering
  - Parametric stochastic neighborhood embedding, Laplacian eigenmaps....
- Use 'smart' classifiers
  - Bayesian neural network, support vector machine, random forest....




---

---

---

---

---

---

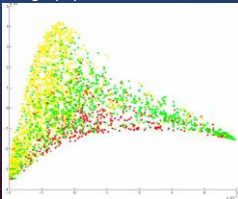
---

---

### RADIOMICS

#### 'SMALL DATA' CHALLENGE: OVERFITTING

Many quantitative computer-extracted image-based features usually describe a single physical characteristic



Unsupervised dimension reduction (Laplacian eigenmaps)

- 1000+ lesions
- 81 image-based features

Cancer  
Benign  
Cystic

AJ Jamieson et al, Med. Phys., 37, 339-351, 2010




---

---

---

---

---

---

---

---

### RADIOMICS

#### 'BIG DATA' CHALLENGE: HARMONIZATION AND INTERPRETATION OF RESULTS

Differences in image acquisition or the population may affect computer-extracted image-based features

- Manufacturer
- Imaging protocol
- Geographic location
  - ❖ Racial differences in disease prevalence and characteristics
- Actual outcome data such as survival may not be available and intermediate alternatives may need to be used
- Need harmonization, especially for less standardized modalities such as MRI
- Correlations amongst the many observations may be vast in number but spurious or unimportant, causality is much harder to assess

N. Gruszkowski et al, Radiology, 253, 661-671, 2009




---

---

---

---

---

---

---

---

---

---

### RADIOMICS

#### REPRODUCIBILITY INVESTIGATIONS



Yesterday morning's lecture by Laurence Court

Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer?

Xenia Fave, Dennis Mackin, Jinzhong Yang, Joy Zhang, David Fried, Pieter Batten, David Followill, Daniel Gomez, A. Kyle Jones, Francesco Stingo, Jonas Fomenot, and Laurence Court

Preliminary investigation into sources of uncertainty in quantitative imaging features

Xenia Fave<sup>1,2,3,4</sup>, Molly Cook<sup>1</sup>, Amy Frederick<sup>1</sup>, Lifei Zhang<sup>1</sup>, Jinzhong Yang<sup>1</sup>, David Fried<sup>1,2</sup>, Francesco Stingo<sup>1</sup>, Laurence Court<sup>1</sup>

X. Fave et al, Med Phys, 42, 6784, 2015

X. Fave et al, Comp. Med. Imaging and Graphics 44, p54-61, 2015




---

---

---

---

---

---

---

---

---

---

### RADIOMICS

#### REPRODUCIBILITY INVESTIGATIONS

RESEARCH ARTICLE  
False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review

Anastasia Chalkidou<sup>1</sup>, Michael J. O'Doherty, Paul K. Marsden

A. Chalkidou et al, PLoS ONE 10(5): e0124165




---

---

---

---

---

---

---

---

---

---





### GENOMICS

#### CHALLENGE: REPLICATION OF RESULTS AND MULTIPLE HYPOTHESIS TESTING

- Challenges here are amplified with respect to radiomics
  - Whole exomes (20-30K genes), whole genomes (3 billion base pairs), with tens of millions of single nucleotide polymorphisms (SNPs) etc.
  - Number of samples ~100<sup>eds</sup>-10,000<sup>nds</sup>

Need correction for multiple-hypotheses testing




---

---

---

---

---

---

---

---

### GENOMICS

#### CHALLENGE: REPLICATION OF RESULTS AND MULTIPLE HYPOTHESIS TESTING

- Want short, reproducible, predictive gene lists (PGLs)

Hypothesis: Genes most important and relevant for control of the malignancy also appear on the list of the most predictive genes

**Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer**

Ein-Dor et al, PNAS, 103, 5923-28, 2006




---

---

---

---

---

---

---

---

### GENOMICS

#### CHALLENGE: REPLICATION OF RESULTS AND MULTIPLE HYPOTHESIS TESTING

- Example: survival prediction in breast cancer

Authors	Technology	Number of patients
Wang et al	Affymetrix	286
van't Veer et al	Rosetta microarrays	96 and 295

Ein-Dor et al, PNAS, 103, 5923-28, 2006

Wang et al, Lancet 365, 671-79, 2005  
van't Veer et al, Nature 415, 530-36, 2002




---

---

---

---

---

---

---

---

**GENOMICS**  
**CHALLENGE: REPLICATION OF RESULTS AND MULTIPLE HYPOTHESIS TESTING**

- Example: survival prediction in breast cancer

Authors	Technology	Number of patients	Number of genes
Wang <i>et al</i>	Affymetrix	286	76
van't Veer <i>et al</i>	Rosetta microarrays	96 and 295	70

Performance was not reproduced, i.e., substantially worse, when tested on different dataset

Only 3 genes in common!

Ein-Dor et al, PNAS, 103, 5923-28, 2006

---

---

---

---

---

---

---

---

---

---

**GENOMICS**  
**CHALLENGE: REPLICATION OF RESULTS AND MULTIPLE HYPOTHESIS TESTING**

- Example: survival prediction in breast cancer

Why these differences?!

- Can **NOT** be explained by only
  - Differences in patient cohorts (such as age)
  - Different microarrays used
  - Different methods of data analysis

Different PGLs obtained from different training sets generated from the same patient cohort typically only have a few genes in common

Ein-Dor et al, PNAS, 103, 5923-28, 2006

---

---

---

---

---

---

---

---

---

---

**GENOMICS**  
**CHALLENGE: REPLICATION....**

- Model for the overlap of top PGLs obtained in replication studies (based on "probably approximately correct" sorting)

Ein-Dor et al, PNAS, 103, 5923-28, 2006

---

---

---

---

---

---

---

---

---

---



**GENOMICS**

**CHALLENGE: REPLICATION OF RESULTS AND MULTIPLE HYPOTHESIS TESTING**

- The good news: P-values are not only highly variable, but the degree of variability is
  - predictable
  - consistent across most types of statistical studies

The prediction intervals for p-value variability can be computed using only

- the p-value of the original study
- the sample size of the replication study relative to that of the original study

Lazzeroni et al, *Molecular Psychiatry*, 19, 1336-1340, 2014




---

---

---

---

---

---

---

---

---

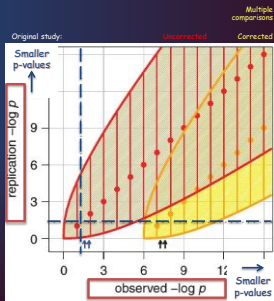
---

**GENOMICS**

**CHALLENGE: REPLICATION....**

- Alzheimer study
- 488,911 SNPs
- 939 individuals
- Original p-value  $\sim 10^{-8}$  for a specific gene association with entorhinal cortical volume

P-value variability prediction for replication study of equal size as original study



SJ Furney et al., *Mol Psychiatry* 16, 1130-38, 2001

Lazzeroni et al, *Molecular Psychiatry*, 19, 1336-40, 2014

---

---

---

---

---

---

---

---

---

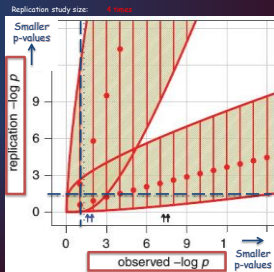
---

**GENOMICS**

**CHALLENGE: REPLICATION....**

- Alzheimer study
- 488,911 SNPs
- 939 individuals
- Original p-value  $\sim 10^{-8}$  for a specific gene association with entorhinal cortical volume

P-value variability prediction for replication study of different sizes than original study



SJ Furney et al., *Mol Psychiatry* 16, 1130-38, 2001

Lazzeroni et al, *Molecular Psychiatry*, 19, 1336-40, 2014

---

---

---

---

---

---

---

---

---

---

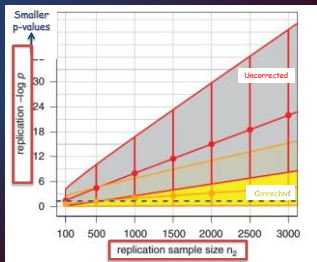


### GENOMICS

#### CHALLENGE: REPLICATION....

- Alzheimer study
- 488,911 SNPs
- 939 individuals
- Original p-value  $\sim 10^{-8}$  for a specific gene association with entorhinal cortical volume

95% confidence intervals for most significant p-value reported in original study



SJ Furney et al., Mol Psychiatry 16, 1130-38, 2001

Lazzeroni et al., Molecular Psychiatry, 19, 1336-40, 2014

---

---

---

---

---

---

---

---

---

---

## REPRODUCIBILITY RADIOGENOMICS




---

---

---

---

---

---

---

---

---

---

### RADIOGENOMICS

Translational Oncology  
www.transonc.com

Volume 7 Number 5 October 2014 pp 556-569 556

**NCI Workshop Report: Clinical and Computational Requirements for Correlating Imaging Phenotypes with Genomics Signatures**

Rivka Colen<sup>1</sup>, Ian Foster<sup>2</sup>, Robert Gatenby<sup>3</sup>, Mary Ellen Giger<sup>4</sup>, Robert Gillies<sup>5</sup>, David Gutman<sup>6</sup>, Matthew Heller<sup>7</sup>, Rajan Jain<sup>8</sup>, Anant Madabhushi<sup>9</sup>, Subhojit Mukherjee<sup>10</sup>, Sandy Ng<sup>11</sup>, Arindam Basu<sup>12</sup>, Joel Saltz<sup>13</sup>, James Tatum<sup>14</sup>, Roshni Verhaak<sup>15</sup> and Gary Whitman<sup>16</sup>\*\*\*




---

---

---

---

---

---

---

---

---

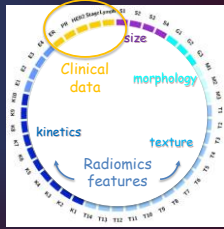
---



### RADIOGENOMICS

#### EXAMPLE EXPLORATORY STUDY

**Investigate:** relation between radiomics, genomics, and combined radiogenomics features and estrogen receptor, progesterone receptor, human epidermal growth factor, tumor stage, and lymph node status



Guo et al, J. Med. Imag., 041007, 2015




---

---

---

---

---

---

---

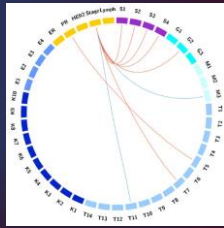
---

### RADIOGENOMICS

#### EXAMPLE EXPLORATORY STUDY

Significant associations between radiomic features and clinical outcomes

**Overall** the prediction performances of genomics alone, radiomics alone, and combined radiogenomics features showed statistically significant correlations with clinical outcomes



Guo et al, J. Med. Imag., 041007, 2015




---

---

---

---

---

---

---

---

## DISCUSSION




---

---

---

---

---

---

---

---

## DISCUSSION

### So...?

Plenty of challenges remain, including computational and statistical aspects. We need

- Harmonization, standards
- Full-disclosure reporting: the way in which predictor variables were chosen and the description of applied significance tests should be as clear as the way in which a dataset was chosen
- Better access to raw data such as publicly available data sets to serve as independent benchmarks

Challenges fuel progress!



---

---

---

---

---

---

---

---

# THE END

Thank you



---

---

---

---

---

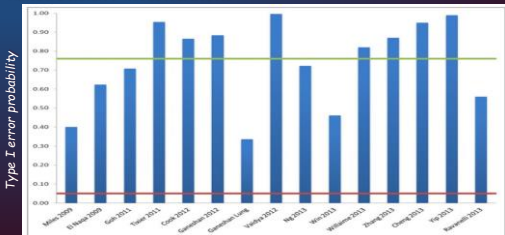
---

---

---

## RADIOMICS

### RERODUCIBILITY INVESTIGATIONS



A. Chalkidou et al. PLoS ONE 10(6): e0124165



---

---

---

---

---

---

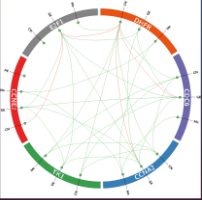
---

---

### GENOMICS

**CHALLENGE:** COMPREHENSIVE PREDICTION OF CANCER GENOMIC INTERACTIONS

**A solution:** Zodiac ([www.compgenome.org/zodiac](http://www.compgenome.org/zodiac))



- Interactions inferred by Bayesian Graphical Model

Zhu et al (2014), Nature Methods

---

---

---

---

---

---

---

---


---

---

### IMAGING GENOMICS

**CHALLENGE:** PROVIDE A BASIS FOR USING NON-INVASIVE IMAGING TECHNIQUES TO INDIRECTLY ASSESS MOLECULAR EVOLUTION OF TUMORS AND THEIR CHANGES UNDER TREATMENT

**PILOT STUDY ON BREAST TUMORS**



Quantitative analysis of TCGA breast MRI cases (Giger lab)

TCGA-Assembler [www.compgenome.org](http://www.compgenome.org)

Analysis of genomics data from corresponding TCGA breast cases (Ji lab)

Imaging Genomics analysis of 91 TCGA breast cancer cases

---

---

---

---

---

---

---

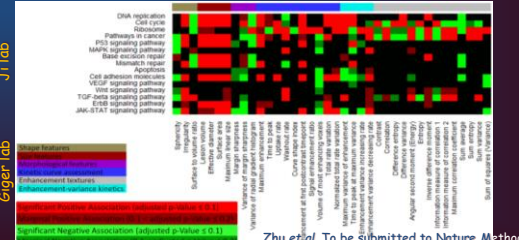
---

---

---

### IMAGING GENOMICS

**PATHWAY TRANSCRIPTIONAL ACTIVITIES ASSOCIATED WITH MRI QUANTITATIVE FEATURES**



Ji lab

Giger lab

Zhu et al. To be submitted to Nature Methods

---

---

---

---

---

---

---

---

---

---

## IMPORTANT STATISTICAL CONCEPTS

STATISTICS GLOSSARY		
Some common statistical concepts and their uses in analyzing experimental results.		
Term	Meaning	Common uses
Standard deviation (s.d.)	The typical difference between each value and the mean value.	Describing how broadly the sample values are distributed. $s.d. = \sqrt{S^2} \quad (S = \text{mean}^2 / (N - 1))$
Standard error of the mean (s.e.m.)	An estimate of how variable the means will be if the experiment is repeated multiple times.	Inferring where the population mean is likely to be, or whether sets of samples are likely to come from the same population. $s.e.m. = s.d. / \sqrt{N}$
Confidence interval (CI, 95%)	With 95% confidence, the population mean will lie in this interval.	To infer where the population mean lies, and to compare two populations. $CI = \text{mean} \pm s.e.m. * t_{(n-1)}$
Independent data	Values from separate experiments of the same type that are not linked.	Testing hypotheses about the population.
Replicate data	Values from experiments where everything is linked as much as possible.	Serves as an internal check on performance of an experiment.
Sampling error	Variation caused by sampling part of a population rather than measuring the whole population.	Can reveal bias in the data (if it is too small) or problems with conduct of the experimental (if it is too large). In binomial distributions (such as life and death cell counts) the expected s.d. is $\sqrt{N * p * (1 - p)}$ ; in Poisson distributions (for example, cells per field) the expected s.d. is $\sqrt{N}$ mean.

*N*: number of independent samples, *t*: the *t*-statistic, *p*: probability.



Di. Vaux, Nature, 492, p180-81, 2012

---

---

---

---

---

---

---

---

---

---