

Grand Challenges in Medical Imaging and Radiomics

**LUNGx Challenge:
Lessons Learned and Key Results**

S.G. Armato III



Disclosures

- Royalties and licensing fees for CAD technology from the University of Chicago
- Consultant, Aduro Biotech, Inc.

Challenges

- Allows comparison of different algorithms
 - all playing by the same rules,

Challenges

- Allows comparison of different algorithms
 - all playing by the same rules,
 - on a common set of images,

Challenges

- Allows comparison of different algorithms
 - all playing by the same rules,
 - on a common set of images,
 - reporting standardized output,

Challenges

- Allows comparison of different algorithms
 - all playing by the same rules,
 - on a common set of images,
 - reporting standardized output,
 - with uniform performance assessment.

Challenges

- Allows comparison of different algorithms
 - all playing by the same rules,
 - on a common set of images,
 - reporting standardized output,
 - with uniform performance assessment.
- Central organization

Challenges

- Allows comparison of different algorithms
 - all playing by the same rules,
 - on a common set of images,
 - reporting standardized output,
 - with uniform performance assessment.
- Central organization
 - time, effort, resources

Challenges

- Allows comparison of different algorithms
 - all playing by the same rules,
 - on a common set of images,
 - reporting standardized output,
 - with uniform performance assessment.
- Central organization
 - time, effort, resources, and crystal ball

Challenges

- Controlled environment
 - database of images
 - training/testing paradigm
 - performance evaluation
 - scoring method
 - reporting structure

Challenges

- Participants deserve much credit
- Friendly competition
- Educational

LUNGx Challenge



Benign



Malignant



LUNGx Challenge

- 2015 SPIE Medical Imaging

LUNGx Challenge

- 2015 SPIE Medical Imaging
- Single-site database of diagnostic CT scans with lung nodules
 - Calibration set: 10 CT scans
 - Test set: 60 CT scans (73 nodules)
 - 12 participants with 15 methods

LUNGx Challenge

- 2015 SPIE Medical Imaging
- Single-site database of diagnostic CT scans with lung nodules
 - Calibration set: 10 CT scans
 - Test set: 60 CT scans (73 nodules)
 - 10 participants with 12 methods

LUNGx Challenge Organizers

- SPIE
- AAPM
- NCI
- University of Chicago
- University of Michigan
- Oak Ridge National Laboratory

Calibration Dataset

- Specifically not a “training set”

Calibration Dataset

- Specifically not a “training set”
- 10 diagnostic CT scans
 - 5 with a benign nodule
 - 5 with a malignant nodule
- Nodule center locations and diagnosis included

Test Dataset

- 60 diagnostic CT scans (73 nodules)
 - 30 with 37 benign nodules
 - 30 with 36 malignant nodules
- Nodule center locations included

Test Dataset

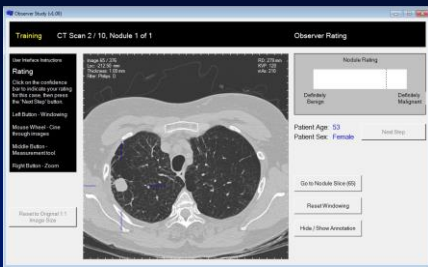
- Participants provided to the organizers
 - a continuous-scale score of nodule malignancy for each of the 73 test set nodules
- Performance assessment managed centrally
 - ROC analysis

Reality Check: Observer Study

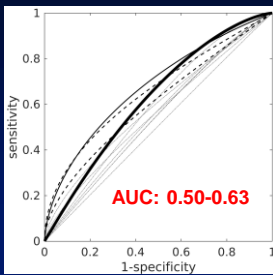
Reality Check: Observer Study

- 6 chest radiologists from two institutions
 - median experience: 14 years (7-41 years)
- Replicate the challenge task

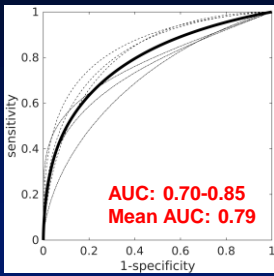
Reality Check: Observer Study



Results: Participants



Results: Observers

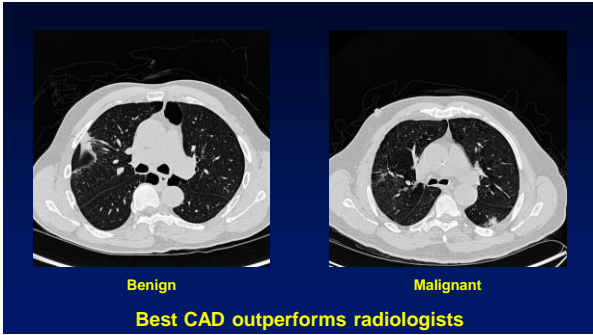


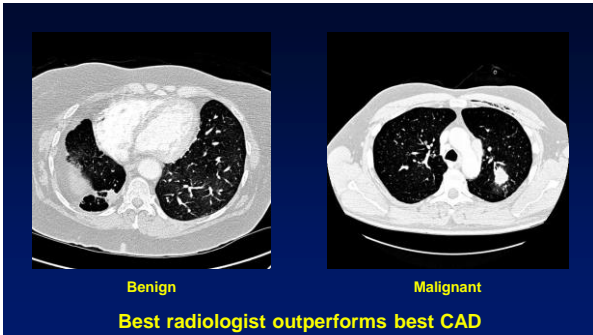
Results

- 3 computerized methods performed better than random guessing
- 3 radiologists performed better than the best computer method

Results

- 3 computerized methods performed better than random guessing
- 3 radiologists performed better than the best computer method
- The significance of these results is lost with correction for multiple comparisons





Journal of
Medical Imaging

MedImaging@SPIEChSTULibrary.org

LUNGx Challenge for computerized lung nodule classification: reflections and lessons learned

Samuel G. Armato, III
 Ljubomir Hadziiski
 Georgia D. Tourassi
 Karen Drukker
 Maryellen L. Giger
 Feng Li
 George Redmond
 Keyvan Farahani
 Justin S. Kirby
 Laurence P. Clarke

Lessons Learned

- Database

Lessons Learned

- Database
 - How many cases? balance effort/cost vs. power
 - Technical parameter consistency
 - slice thickness
 - contrast enhanced
 - Lesion characteristics consistency
 - size
 - histology

Lessons Learned

- Database
- Instructions

Lessons Learned

- Database
- Instructions
 - Acceptable level of human input
 - Appropriate use of test set
 - Don't assume

Lessons Learned

- Database
- Instructions
- Clinical data

Lessons Learned

- Database
- Instructions
- Clinical data
 - limit to DICOM header (age, gender)?

Lessons Learned

- Database
- Instructions
- Clinical data
- **Commitment by participants**

Lessons Learned

- Database
- Instructions
- Clinical data
- **Commitment by participants**
 - anonymity
 - ability to withdraw

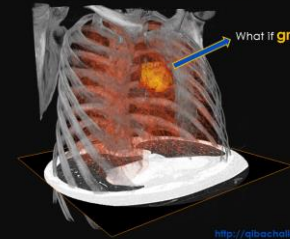
Lessons Learned

- Database
- Instructions
- Clinical data
- Commitment by participants
- **Epilogue**

Lessons Learned

- Database
- Instructions
- Clinical data
- Commitment by participants
- Epilogue
 - image data persistence
 - publish?

CT Virtual Clinical Trial Grand Challenge



What if ground truth is known?

Test the precision and accuracy of your algorithm to quantify lesion volume.

<http://qibchallenges.dukeup.edu>





- SPIE-AAPM-NCI PROSTATEx Challenge 2017



