2013 Report from the Volumetric CT Technical Committee of the Quantitative Imaging Biomarkers Alliance

Gregory Goldmacher, MD, PhD, Samuel Armato III, PhD, Maria Athelogou, PhD, Andrew Buckler, MS, Nicholas Petrick, PhD, Binsheng Zhao, DSc, Lawrence Schwartz, MD

Quantitative Imaging Biomarkers Alliance RSNA

Background And Previous Work

Establish 3D CT Volumetry as a Surrogate Endpoint Biomarker for Tumor Response

Specific Aims

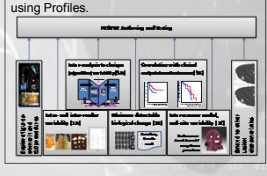
To develop methods and processes for accurate and reproducible measurements of anatomic structures and masses.

Stakeholders

FDA, NIST, NCI, FNIH, ACRIN, manufacturers and developers, CROs, PINTAD, and academic centers.

Approach

Process map for qualifying volumetric image analysis using computed tomography, showing relationship to standardization of the biomarker's measurement and interpretation using Profiles.



Where We Go From Here

- Advanced Disease:
  1. Complete our project in lung cancer by developing methods for quantifying its metastases to the liver and lymphatic system; and
  2. Expand these results to a broader set of indications that also produce solid metastatic lesions in the liver and the lymphatic system. These tissues are common sites of disease in colorectal, esophageal, renal, breast, stomach, pancreatic, melanoma, etc.
- Lung Nodule Assessment in CT Screening

Federal Funding Acknowledgements

Various QIBA projects and activities have been funded in whole or in part with Federal funds from the National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN268201000050C.

Various QIBA projects and activities have been supported in part with Federal funds from the National Institute of Standards and Technology, Department of Commerce, under Cooperative Agreement No. 70NANB10H223 (which partially funds the QI-Bench program).

What We've Accomplished

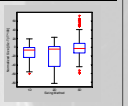
- Completed and reported charter studies (see below).
- Expanded studies to characterize performance under an increasing range of settings.
- Written, refined, and released Profile through Public Comment and Field Test processes.

Sizing of Synthetic Spherical and Non-Spherical Lung Nodules (aka "1A")

Aim

To estimate bias/variance of radiologists estimating the size of synthetic nodules

Results



- Box and whisker plot of reader size estimates for the 1D, 2D and 3D sizing methods (all phantom nodules included).

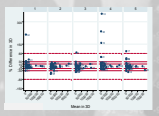
Summary

- Overall, 3D method provided low bias estimates of nodule volumes
- 3D method applied to thin slice data provided low bias & low variance estimates

Minimum Detectable Change in Clinical Trial Workflows (aka "1B")

Hypothesis The minimal detectable change in tumor size will be smaller when using a side by side ("clinical trial workflow") review setting than when using timepoints presented in random order.

Results



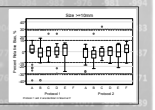
Discussion

1. Measurement variability is considerably reduced when using the locked, sequential read approach compared to randomized timepoint reads
2. Should inform the QIBA profile as to "best practices" for clinical trials

Evaluating inter CT scanner effects in clinician sizing of phantom nodules (aka "1C")

Aim: characterize accuracy and precision in reader measurements of volumes of six phantom nodules collected on six scanners.

Results



Conclusion

- Relative bias is within a tolerance of 15%.
- Scanner equivalence is found only for the larger synthetic lesions (10 mm and 20 mm). This finding confirms the lesion sizing guidance (10 mm and up) in QIBA CT imaging profile.
- Equivalence of the protocols supports the imaging protocol as used by ACRIN Trial 6678.

2013 Progress

Inter-algorithm Performance Investigation Using Phantom Data

QIBA 3A Group

Challenge Definition: estimate absolute volumes using **CT phantom data.** Explicitly indicate descriptive statistics: **bias, variance.**

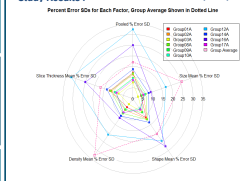
Null hypothesis: analysis software model does not have a significant effect on bias and variance.

FDA, M. A Gavrielides et al., "A resource for the assessment of lung nodule size estimation methods: database of thoracic CT scans of an anthropomorphic phantom", Optics Express, vol. 18, n.14, pp. 15244-15255, 2010.

Phantom data produced by the FDA, and available on the QI-Bench website. Support provided by NIST.

Method: The study was organized as a public challenge. Computer Tomography (CT) Scans of synthetic lung tumors in anthropomorphic phantoms were acquired by the Food and Drug Administration (FDA). Their physical measurement values were used as ground truth in order to calculate the algorithm measurement bias and variability. Tumors are varied in size, shape, and density. The participants downloaded the images as well as coordinates and bounding boxes for each tumor. Descriptive statistics and Analysis of Variance (ANOVA) were used to test the characteristics of the phantoms and their software-based measurements in terms of volume bias.

Study Results : 10 participants who measured 408 nodules



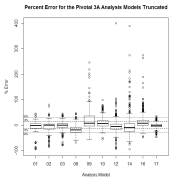


Figure 1 Percent Error for all Participants (without the Reference markups): Using only the nodules that is met QIBA CT profile, the standard deviation from pooled data for all 10 participants are shown by the dotted pink polygon with a pink dotted polygon. The pooled standard deviation of each 10 participant is shown by the different colors with a polygon.

Figure 2 Box-whisker plot representing the distribution of the percent error in volume measurements. The mid-bold line indicates the median. The upper and lower lines of box represents 25% and 75% tile in the percent errors. The thicker dashed lines represent ±15%, and the smaller dotted lines show the location of ±30%.

Validation of volumetric CT as a better imaging biomarker for predicting patient survival

QIBA 3B Group

Specific Aims

1. To explore variability in measuring change in tumor volume (uni- and bi-dimensional as well)
2. To correlate early responses assessed by the volumetric, uni-dimensional and bi-dimensional measurement techniques with patient survival

Materials and Methods

- 560 patients enrolled to a multicenter Phase II / III clinical trial in advanced colorectal cancer treated with a targeted therapy were included in this study.
- Target lesions were measured on baseline and 6-week follow-up study using in-house lesion segmentation algorithms developed for solid tumors.

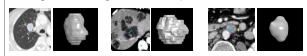


Figure 1. Computer-aided tumor segmentation and quantification

SA #1

- A subset of 30 patients' baseline and 6-wk follow-up CT scan data were analyzed.
- Three radiologists participated in the variability study
- Per RECIST each radiologist independently selected and, with the help of computer software, measured targeted lesions on the two scans.
- Variability was analyzed using linear mixed effects models and Bland-Altman method.

SA #2

- 560 patients' CT data (on average, 7 scan time-points per patient) were analyzed.
- Tumor responses to therapy were assessed using RECIST (uni), WHO (bi) and a new volumetric technique.
- The overall survival will be compared using Kaplan Meier graphs and Cox regression models to estimate the hazard ratios for the various categories of response.

Results

SA #1

	Unidimensional			Bidimensional			Volume		
	Relative change @ 6-week			Relative change @ 6-week			Relative change @ 6-week		
	Mean	95% Limits of Agreement		Mean	95% Limits of Agreement		Mean	95% Limits of Agreement	
R1 vs R2	-3%	-18%	11%	-4%	-30%	23%	-2%	-38%	34%
R1 vs R3	-1%	-16%	13%	0%	-27%	27%	0%	-34%	34%
R2 vs R3	2%	-13%	17%	4%	-25%	32%	2%	-26%	29%

Table 1. Variability in relative change of total tumor burden measured unidimensionally, bidimensionally and volumetrically (1,2)

SA #2

Clinical correlation analyses are being evaluated.

1. Zhao B, Lee S, Lee HJ, et al, Relationship of variability in tumor measurement and response. ASCO June 1-5, 2012, Chicago.
2. Zhao B, Lee S, Qi J, et al, Inter- and intra-reader variability in assessing change of total tumor burden. CARS June 26-29, 2013, Heidelberg, Germany.

Lung Nodule Assessment in CT Screening: The "Small Nodule" Group

Background:

• The QIBA CT Volumetry Technical Committee has completed its effort to define a process for optimal quantification of advanced lung cancer masses.

• The task now is to address analogous issues in the use of quantitative imaging to segregate clinically significant from insignificant lung nodules in the CT-based lung cancer screening process.

Purpose:

• To define evidence-based consensus standards and processes for CT imaging to allow for reproducible nodule characterization and quantification of biologically meaningful longitudinal volume changes with an acceptable range of variance across vendor platforms

• To develop standardized methods for performing repeatable volume measurements on CT images of lung nodules in the setting of lung cancer screening and post-screening surveillance

Rationale:

• CT provides an effective means of detecting and monitoring lung nodules and can lead to a reduction in mortality in individuals at high risk for lung cancer.

• Size quantification on serial imaging is helpful to evaluate whether a nodule is benign or malignant.

• Quantification of nodule volume (volumetry) could solve some of the limitations of current diameter-based measurement metrics.

The Challenge:

• CT screening presents a challenge for the development of an optimized protocol in the need to balance the risks and harms in an asymptomatic population: lowest dose possible vs. reduced image quality that would limit the ability to detect small nodules.

• The extent to which the increased noise associated with a lower dose affects radiologists' ability to accurately measure small nodules is being mitigated by evolving technology.

• Any protocol will represent a compromise between these competing needs in the screening setting.

The QIBA Profile:

• will make claims about the confidence with which changes in lung nodule volume can be measured under a set of defined image acquisition, processing, and analysis conditions.

• will provide specifications designed to allow the achievement of targeted levels of clinical performance in identified settings within the screening and surveillance context.

• will be intended for all stakeholders invested in lung cancer screening, including designers of CT image-acquisition protocols, those involved in the procurement of new CT equipment, software and device manufacturers, pharmaceutical companies, oncologists, clinicians engaged in the screening process and in clinical trials, those who obtain quantitative measurements on CT images, and those who make decisions based on quantitative image measurements.

Future Directions

Inter-algorithm Performance Investigation Using Clinical Data

QIBA 3A Group

Challenge Definition: Estimate absolute volumes in **clinical CT data.** Explicitly indicate descriptive statistics: **bias, variance.**

Null hypothesis: analysis software model does not have a significant effect on the bias and variance.

Study Status:
- CT – clinical data is at QI-BENCH Website (www.qi-bench.org)

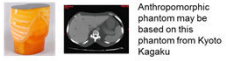
- result submission from 11 participating sites with 12 algorithms

- results are uploaded on QI-BENCH Website

Nest Step:
- statistical analysis of the study

Accuracy and Precision of Liver Lesion Volume Sizing Tools

- Aim: Assess measurement performance of lesion sizing tools in estimating the volume of liver lesions
- Methods
  - Anthropomorphic Phantom



Anthropomorphic phantom may be based on this phantom from Kyoto Kagaku

  - Synthetic Nodules

  - CT scans of hyper- and hypo-dense liver lesions of various shapes and sizes
- Study Design
  - Designed as public challenge
    - Multiple participants download images with known liver lesions
    - Apply their automated or semi-automated volume sizing tool
    - Upload volume & lesion boundary estimates
- Analysis
  - Accuracy: statistical comparison of biases
  - Precision: statistical comparison of reproducibility among tests
  - ANOVA analysis to identify important imaging & lesion factors affecting volume measurements

Test / re-test Reproducibility of CT Volumetry in an Animal Model of Liver Tumors

Background:

To define the minimal detectable change in CT volumetry of liver metastases, we need information about test/re-test reproducibility. Phantom measurements cannot capture key variables that affect liver lesion measurements, including the delivery of contrast through the liver's dual blood supply. A test/re-test study needs to be done in a biological system. Getting such data in humans may be difficult because of ethics/safety concerns. A pig model of liver metastases has been developed, and will be used to characterize test/re-test variability.

Project Plan:

- Modified fibroblast cell line samples implanted into liver of immunosuppressed pigs (3-4 lesions/pig).
- Tumors grow for ~2 months, with occasional size check by ultrasound, until ~2 cm diameter.
- Pig anesthetized, and CT scans performed at several intervals after injection of IV contrast.
- After ~12 h to allow contrast washout, scans repeated, taking care to keep all contrast injection parameters (volume, rate, scan timing) the same.
- Process repeated once more, for a total of 3 sets of scans obtained over approximately 24 hours.

Other current and future activities

- Development of compliance requirements and testing process to support the published profile
  - Actions required of each "actor" (scanner, technologist, radiologist reconstruction software) involved in the scanning and analysis process
  - Performance standards for results measured after the process has been followed
- Assessment of the existing CT volumetry profile
  - Has it been implemented, and how widely?
  - Technical validation – how does it function when applied as written?
- Collaboration with groups working on clinical validation of volumetry in predicting patient outcomes and clinical trial results

Various QIBA projects and activities have been funded in whole or in part with Federal funds from the National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN268201000050C.