Technical Performance of CT Measurement of Changes in Tumor Volume

QIBA CT Volumetry Technical Committee*

Quantitative Imaging Biomarkers Alliance — RSNA

Overview

Specific Aims

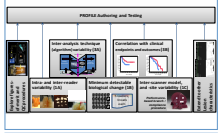
To develop methods and processes for accurate and reproducible measurements of anatomic structures and masses.

Stakeholders

FDA, NIST, NCI, FNIH, ACRIN, manufacturers and developers, CROs, PIG, 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.



What We've Accomplished

- Charter studies have reported results.
- Expanded studies to characterize technical performance under an increasing range of settings and to establish clinical performance.
- Written, refined, and released Profile and Protocol documents through Public Comment and Field Test processes.

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).

*Author list and PDF copy of this poster are available online at:
http://qibawiki.rsna.org/index.php?title=QIBA_posters_and_plenary_session_presentation_from_RSNA_2012_Annual_Meeting

Determining Minimum Detectable Change in Patient Datasets: Results from Clinical Trial Workflow Conditions

Hypothesis

That the minimal detectable change in tumor size - using measured tumor volumes made by radiologists on thin section CT images - will be smaller when using a side by side ("clinical trial workflow") review setting than when using an independent review setting (previous study)

Methods

1. CT images of 52 lesions (1 per patient) were utilized; each with two time points. 32 were "coffee break" or "no change" condition cases, while 20 cases had change; the latter cases served as distracters so readers would not expect all cases to be "no change" cases.
2. 5 readers contoured the boundary of each lesion completely in 3D in a sequential, locked reading paradigm. The first time point was contoured and then locked before the second time point was displayed. The second time point was displayed side-by-side to the first case so direct comparisons were allowed. For each case, boundaries were used to calculate volume
3. Calculate percent difference between time points for "no change" cases. Determine 95% confidence interval of "no change"; Compare results to previous study (reported RSNA 2011) where a different reading paradigm of independent reads were performed.

Results

	Independent Reads [95% CI]	Locked, Sequential Reads [95% CI]
%Volume Change	[-4.0%, 55%]	[-2.2, 17%]
% Diameter Change	[-0.3%, 12%]	[0.3%, 4.7%]

Table 1 – 95% confidence intervals for volume and diameter measurements made in "no change" cases for both Independent reading (previous study) and Locked, Sequential readings (this study)

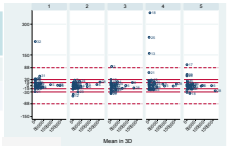


Figure 1- Graph showing % Difference in Volumes between time points as a function of lesion volume; each column represents a different reader – data labels are case numbers

Discussion

1. Measurement variability is considerably reduced when using the locked, sequential read approach compared to independent 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

Aim

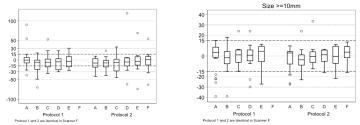
Characterize accuracy and precision in reader measurements of volumes of six phantom nodules in CT imagery collected on six scanners.

Methods

1.Develop and apply two arms of the imaging protocol
- Imaging protocol as used by ACRIN Trial 6678 (Lung Cancer: Evaluation of Treatment Response with PET)
- Quality-based arm targeting constant noise and resolution across scanners

2.Image anthropomorphic phantom at six imaging sites on six scanners: a Siemens Sensation 64, Toshiba Aquilion 64, Philips Brilliance 64, two Philips Brilliance 16s, and a GE VCT 64.
3.Thorax phantom has 6 spherical and 6 spiculated nodules, 6 in each 'lung'. The sizes are 5 mm, 10 mm and 20 mm (volume of an equivalent sphere.) We read and analyze six of the nodules.
4.Readers measure semi-automated 3-D volume. 1-D & 2-D sizings are derived.
5.Analysis – accuracy, precision and variance between scanners; protocol arms; reader s; & nodule size and shape.

Summary Data: Percent Relative Bias in volume by Scanner and CT protocol



Conclusions:

- Relative bias in pooling the 6 nodules is within a tolerance of 15%.
- In an equivalence t-test applied to20 mm each of the 6 nodules, scanner equivalence is found only for the larger synthetic lesions (10 mm and). 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.

Inter-algorithm Performance Investigation Study 3A Group

Phantom data, FDA, NIST, QI-

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.

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

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

Investigation 1:
Pilot and pivotal study are finished

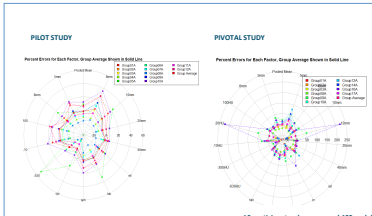


Fig. 1: Radial plot showing comparative performance on the selected descriptive statistics as well as mean of absolute percent errors.

Development of Assessment and Predictive Metrics for Quantitative Imaging in Chest CT

Aims

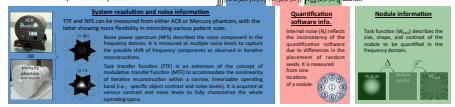
Develop and evaluate a metric (estimability index, e') capable of modeling/predicting the performance of chest CT volume quantification.

Methods

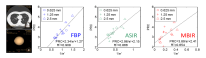
1.Develop a framework for prediction of volume quantification precision

Estimability index (e') is a **predictive metric**, of quantitative precision:

$$e' = \frac{\iint [TTF^2(u,v;\alpha,C,M)]\,W_{obj}(u,v,w')\,du\,dv\,dw'}{\iint [NPS(u,v,w')]\,W_{obj}(u,v,w')\,du\,dv\,dw'}$$



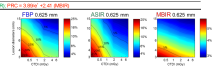
2.Validate/calibrate e' against the precision measured from an anthropomorphic phantom in terms of percent repeatability coefficient (PRC)



3.Establish a step-by-step instruction for e' -based prediction of precision

Phase 1	Step 1	Use phantom that contains cylindrical inserts of various attenuations for TTF measurements and an uniform region for NPS measurements. Scan and reconstruct the phantom with a range of parameters (mAs, slice thicknesses, kernel, etc.) to compute a library of 3D TTF and NPS values that characterize the entire operating space
Phase 2	Step 2	Model W_{obj} according to the size, shape, and contrast of the nodule being assessed *Model N according to the quantification software being assessed*
	Step 3	Interpolate a 3D TTF from the library built in Step 1 with respect to the nodule's contrast and the image noise of the protocol (imaging and reconstruction parameters) being assessed. This is important for protocols involving iterative reconstructions. Interpolate a 3D NPS from the library with respect to the image noise of the protocol
	Step 4	Incorporate TTF, NPS, W_{obj} and N into calculating e' for the aforementioned combination of nodule, segmentation software, and protocol
	Step 5	Relate e' to PRC: PRC = 2.34e' +1.27 (RBP); PRC = 2.90e' +2.15 (ASR); PRC = 3.80e' +2.45 (MBR)

4.Predict quantitative precision from e' for recommending guidelines for compliance of quantification techniques



* For more details, please contact baiyu.chen@duke.edu

Current Status

Standardize the calculation of e' and recommend guidelines for compliance of quantification techniques

Quantifying Volumes of Part-Solid Nodules

Aim

Extend characterization of nodule measurement performance to the part-solid case in a phantom study. Primary endpoints include bias and variability relative to known nodule volume, with covariates including dose, slice thickness, nodule shape, size, and mean CT density.

Methods

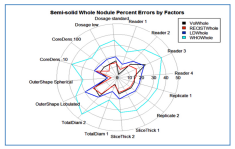
Part Solid and Solid Nodules
Shape, Size (Whole/Core), Density (Outer,Core)
Spherical, 10mm/5mm, -630 HU/-10HU
Spherical, 10mm/5mm, -630 HU/100HU
Lobular, 10mm/5mm, -630 HU/-10HU
Lobular, 10mm/5mm, -630 HU/100HU
Spherical, 20mm/10mm, -630 HU/-10HU
Spherical, 20mm/10mm, -630 HU/100HU
Lobular, 20mm/10mm, -630 HU/-10HU
Lobular, 20mm/10mm, -630 HU/100HU
Spherical, 12mm/NA, 100HU/NA
Spherical, 10mm/NA, 100HU/NA
Spherical, 8mm/NA, 100HU/NA
Spherical, 5mm/NA, 100HU/NA

Acquisition Protocol
QIBA Profile. Computed Tomography: Change Measurements in the Volumes of Solid Tumors, Version 2.0
Siemens Sensation 64
Slice Thickness: 1mm, 2mm
Dose: 40mAs, 200mAs

Reading Protocol
Four Radiologists, 2 reads each
Semi-automated (SA), seed-based, INTIO research software
Data output (ALL based on the 3D segmentation):
Nodule Volume
SA-RECIST: Axial plane diameter
SA-LD: Longest diameter in an orthogonal plane
SA-WHO: Area based on WHO criteria

Significant Results – Whole Nodule Measurements

- Absolute Bias of Whole PS Volume > SA-RECIST
  - PS Volume Absolute Bias ~ 20.38
  - SA-RECIST Absolute Bias ~ 15.64
  - SA-RECIST taken from volume segmentation
  - Spherical Nodule bias
- Significant Covariates of Volume and SA-RECIST
  - Nodule Diameter
  - Nodule Outer Shape
  - Reader (R)
- Significant Covariates of SA-RECIST only
  - Nodule Core Density
- ICC (95% CI)
  - Whole Volume Measures – Reader Average
    - 0.955 (xxx, 1.00)
  - Absolute bias from Volume – Reader Average
    - 0.302 (xxx, xxx)
- No significant effect of dose



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

Specific Aims

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

Materials and Methods

- We used an image dataset of 560 patients enrolled in a multicenter Phase II / III clinical trial of advanced colorectal cancer and treated with a targeted therapy.
- Target lesions were measured on baseline and follow-up scan time-points using in-house lesion segmentation algorithms developed for solid tumors (Fig.1).

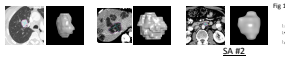


Fig 1. Computer-aided tumor measurements

SA #1

- A subset of 30 patients' baseline, 6-wk and 12-wk follow-up CT scans were analyzed.
- Three radiologists participated in the inter-reader variability study, and two of them also participated in an intra-reader study.
- Each radiologist independently measured target lesions on three 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 measured.
- 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[1]

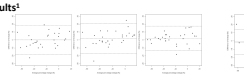


Fig 2. Inter-reader (a-c) and intra-reader (d) variability in uni relative changes at 12-wk from baseline

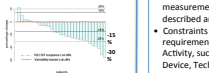


Fig 3. Waterfall plot – relative change at 6-wk

- Inter- and intra-reader variability in change of uni-dimensional tumor burden were similar.
- 95% limits of agreement on uni-dimensional tumor burden were all within (-15%, 15%).

Work-in-Progress

Variability in measuring change in tumor volume is under analysis, as is the correlation analysis of SA #2.

1. Zhao B, Lee S, Lee HJ, et al, Relationship of variability in tumor measurement and response. ASCO June 1-5, 2012, Chicago.

QIBA Profile: CT Tumor Volume Change (CTV-1)

Aim

Develop a technical specification document ("QIBA Profile") describing the capabilities, performance and procedures for equipment and staff to achieve a stated "Profile Claim".

The Claim of this Profile is:

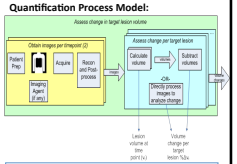
A measured volume change of more than 30% for a tumor provides at least a 95% probability that there is a true volume change; i.e.
P(true volume change > 0% | measured volume change >30%) > 95%.

The scope of the claim is limited to tumors that are measureable (i.e., tumor margins are sufficiently conspicuous and geometrically simple enough to be recognized on all images in both scans), and the longest in-plane diameter of the tumor is 10 mm or greater.

Deliverables

Download Profile document here:
http://rsna.org/QIBA_Protocols_and_Profiles.aspx

Quantification Process Model:



Method

- The document was drafted by the QIBA Technical Committee on Quantifying Volumes with X-ray Computed Tomography which includes radiologists, physicists, statisticians, cancer researchers, imaging vendors and pharmaceutical clinical trialists.
- The specifications and claims are based on groundwork such as that shown to the left.
- The draft Profile was published for Public Comment in late-2011 and the comments resolved in 2012
- The Profile is now available for consideration and trial implementation by equipment vendors and clinical trialists.

Content

- Four Activities in the quantification process are profiled: Subject Handling, Image Acquisition, Image Reconstruction and Image Analysis.
- In each Activity, details considered to impact measurement quality and reproducibility are described and constrained.
- Constraints are stated as compliance requirements on the Actors that perform the Activity, such as the Radiologist, Acquisition Device, Technologist & Image Analysis Tool.
- Details include patient positioning, contrast use, characteristics of acquisition and reconstruction protocols, analysis tool features and performance.
- The Profile includes methods to record specific parameter values that result in compliant performance on specific equipment models.