CT Lung Nodule Volume: Comparison 2D and 3D Analysis

Philip F. Judy
and
Francine L. Jacobson

Department of Radiology
Brigham and Women's Hospital
and
Harvard Medical School
Boston, MA

judy@bwh.harvard.edu

Introduction

If lung cancer screening using CT is to succeed, then procedures must exist to account for the differences in CT equipment and these analysis procedures. We are proposing an approach to evaluate the differences through the use of a lung nodule phantom. We demonstrate its use by comparing several procedures that estimate the size of nodules from the same CT image set. Specifically, validity of CT size measurements of acrylic balls in lung simulating materials measured by 2D and 3D computer analysis procedures were compared. Computer analysis procedures have the potential of being more reliable that human measurements (Jacobson-2000)

Jacobson-2000, this RSNA, Paper 187, page 243, Validity of Objective and Subjective Measurements of Small Pulmonary Nodules from CT images

Background

Clinical trials are evaluating early detection of lung cancer using CT (Sone-1998, Henschke-1999). The primary indication of malignancy of nodules of diameters less than 5 mm is growth measured with sequential CT scans (Zhao-1999, Yankelevitz-1999). Automated procedures to evaluate have been developed (Brown-2000, Armato-2000, Reeves-2000, Campostrini-2000).

Armato-2000, this RSNA, Paper 186, page 243, A Computer-aided Diagnostic Methods for the Detection of Lung Nodules in CT Scans.

Brown-2000, this RSNA, Paper 185, page 243, Computer-aided Method of Lung Micronodule Detection in CT.

Campostrini-2000, this RSNA, Paper 185, page 263, Estimating Irregular Volumes by Means of the CT-Based "Density Mask" Program: A Promising Tool to Evaluate Lung Cancer Response to Therapy.

Henschke CI, McCauley DI, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS, Libby DM, Pasmantier MW, Koizumi J, Altorki NK, Smith JP. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet, 1999 July 10;354(9173):99-105.

Reeves-2000, this RSNA, Paper 188, page 243, Analysis of Small Pulmonary Nodules without Explicit Segmentation of CT Images.

Sone S, Takashima S, Li F, Yang Z, Honda T, Maruyama Y, Hasegawa M, Yamanda T, Kubo K, Hanamura K, Asakura K. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet. 1998 Apr 25;351(9111):1242-5.

Yankelevitz DF, Gupta R, Zhao B, Henschke CL. Small Pulomary Nodules: Evaluation with Repeat CT – Preliminary Experience. Radiology 1999; 212: 561-566.

Zhao B, Reeves A, Yankelevitz D, Henschke C. Two-dimensional multi-criterion segmentation of pulmonary nodules on helical CT images. Med Phys 1999; 26:889-895.

Lung Nodule Phantom

The lung nodule phantom (Figure 1) contains 15 acrylic balls placed in a 15 cm disk of simulating lung material (-700 HU, The Phantom Laboratory). The nominal diameters of the balls are 2 to 8 mm.

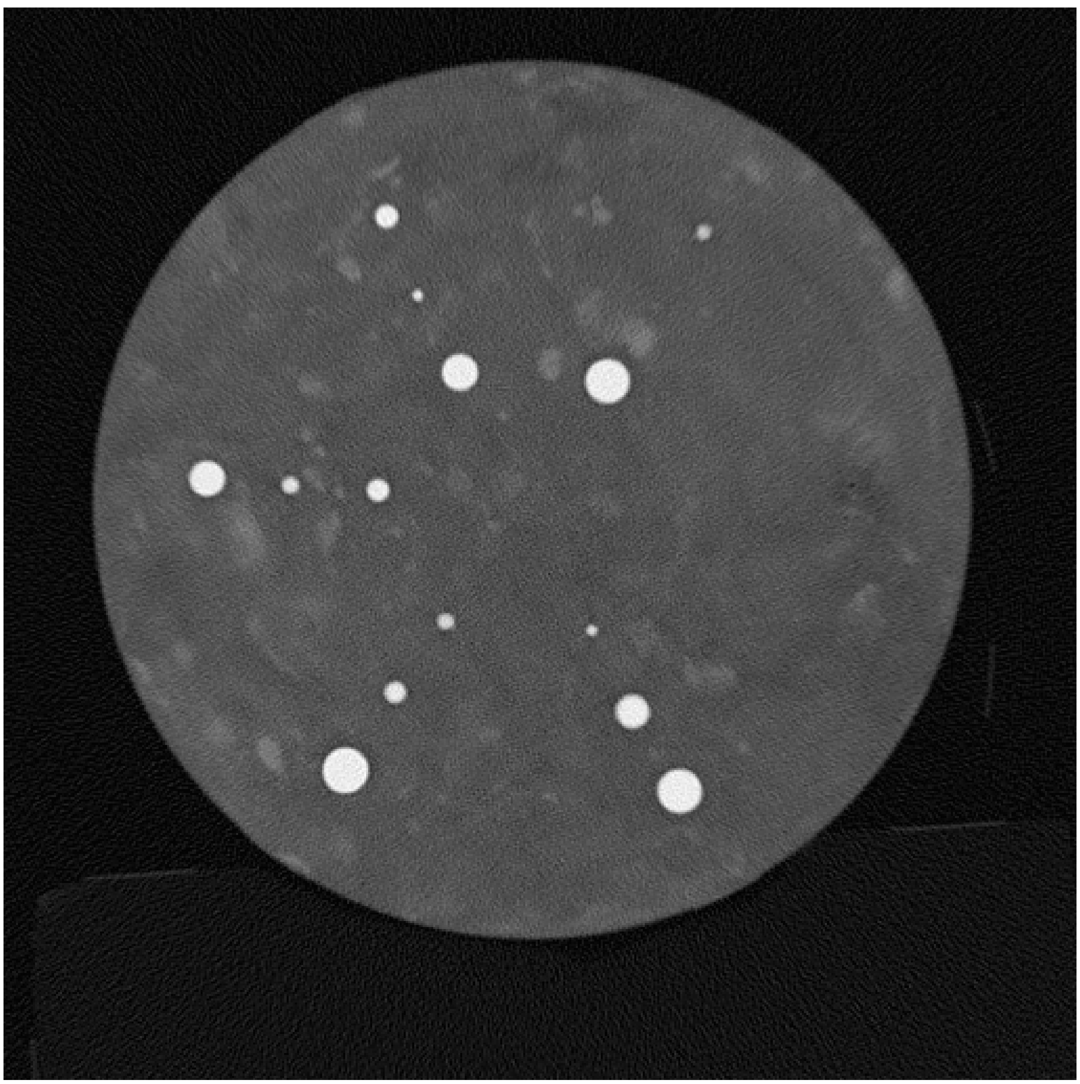


Figure 1 --- CT of Lung Nodule Phantom

Lung Nodule Phantom

The acrylic balls were weighed before fabrication of the phantom. Reference volumes and reference diameters (Table 1) of the balls were calculated assuming spherical balls and acrylic density of 1.19 g/cm^3.

Nominal Diameter (mm)	Weight (mg)	Reference Volume (cubic mm)	Reference Diameter (mm)
2	5.5	4.62	2.07
3.175	21	17.25	3.21
4	41	34.1	4.02
6	159	133.8	6.35
8	305	256.2	7.88

Table 1 - Mean size of acrylic balls

Lung Nodule Phantom

Lung nodule phantom placed on a patient during a lung cancer screening exam. Reconstructed section thickness was 8 mm. The lung nodule phantom carried a TLD to estimate the radiation exposure (0.900 mR) to the breast.

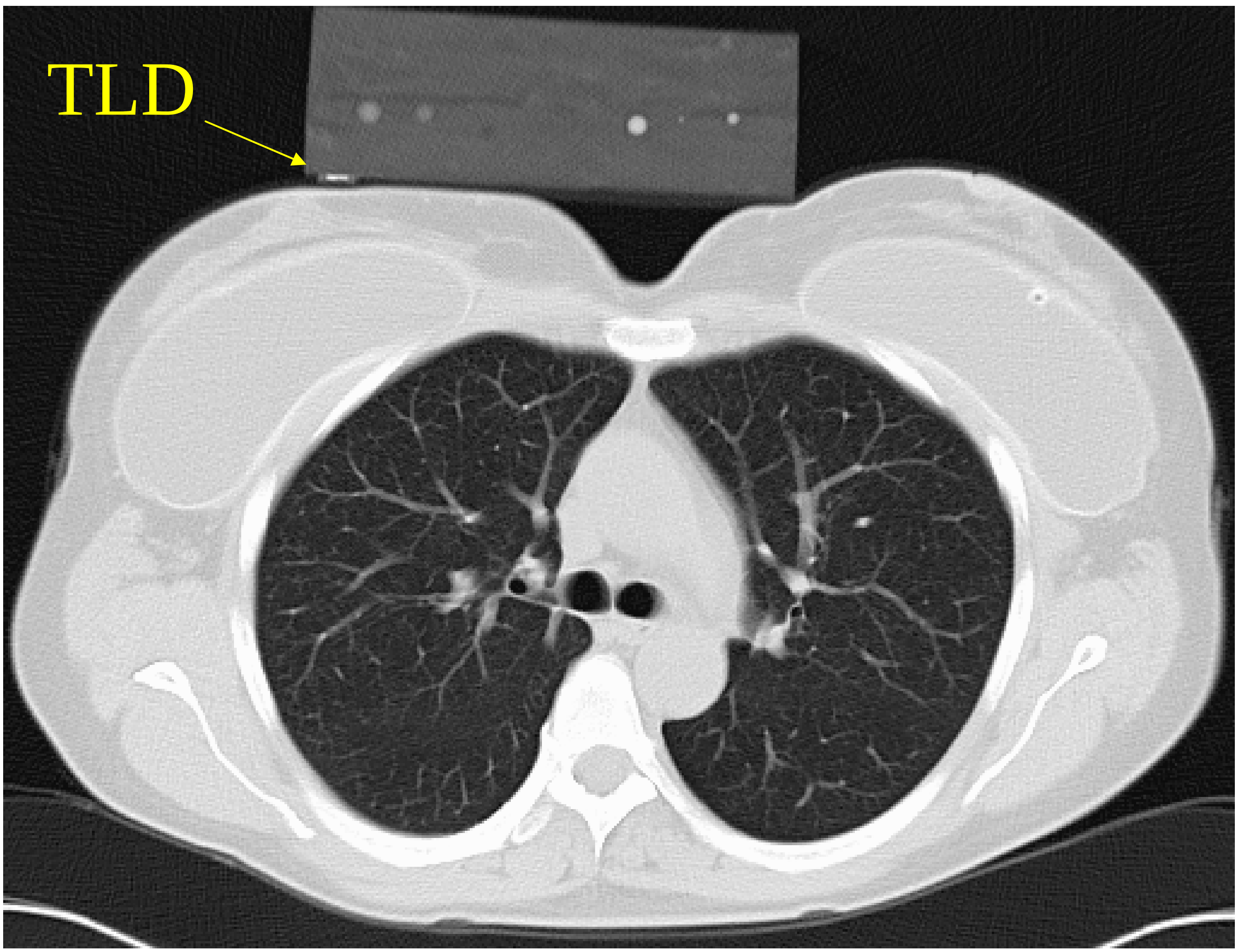


Figure 2 --CT of Chest, TLD, and Lung Nodule Phantom

Lung Nodule Phantom: CT Scan Technique

- Figure 1 is CT section of the lung nodule phantom through the plane of acrylic balls.
- Multi-detector CT scanning of the lung nodule phantom was performed with Siemens Volume Zoom CT scanner using 1.0-mm collimation with pitch of 4, inter-section distance of 0.5mm, and 1.25 mm slice sensitivity profile width. The pixel size was 0.3574 mm. The scan time was 0.75 sec using a tube current of 67 mA and tube potential of 140 kV. Reconstruction filter was designated B70s.

Reference Standards

- A <u>reference standard</u> is procedure that provides an estimate of size that is more widely accepted as accurate than the procedure under study. A reference standard that is the best is called a "gold standard." Our gold standard is based on the weight of the acrylic balls.
- <u>Image reference standards</u> are calculations that use the image data to estimate size that should be "reasonably" accurate.
- Since the image data is degraded, an <u>image reference standard</u> should less accurate than a <u>gold standard</u>. (There is no guarantee that a procedure exists to accurately estimate the size of a feature. The image data may be so degraded by unsharpness that information is lost)

Threshold Procedure

- A simple thresholding procedure is a reasonable image reference standard for the lung nodule phantom image, because the background is uniform.
- Thresholding means that the area of each nodule in each section was determined from the number of pixels contained in each nodule feature that were above a specified threshold level. The threshold value was identical for all images in each image set. This procedure produced area estimates as function of axial location (Figure 3)
- Maximum area and axial extent procedures should be the same for the geometric cross-sections of a spherical ball. In Figure 3, axial extent agrees with gold standard, but maximum area doesn't.

Different threshold results are possible.

Figure 3 demonstrates that image reference areas for each section are less than gold standard areas. The axial extent of the image features is equal to gold standard diameter. Consequently, three different image reference diameters were determined from an analysis of areas as function of axial location of the section.

1. Sum of areas, i.e., the volume of the nodule making no assumptions about the shape of the nodule
2. Maximum area from the parabolic fit (area versus axial index)
3. Axial extent from the parabolic fit (area versus axial index)

Threshold areas

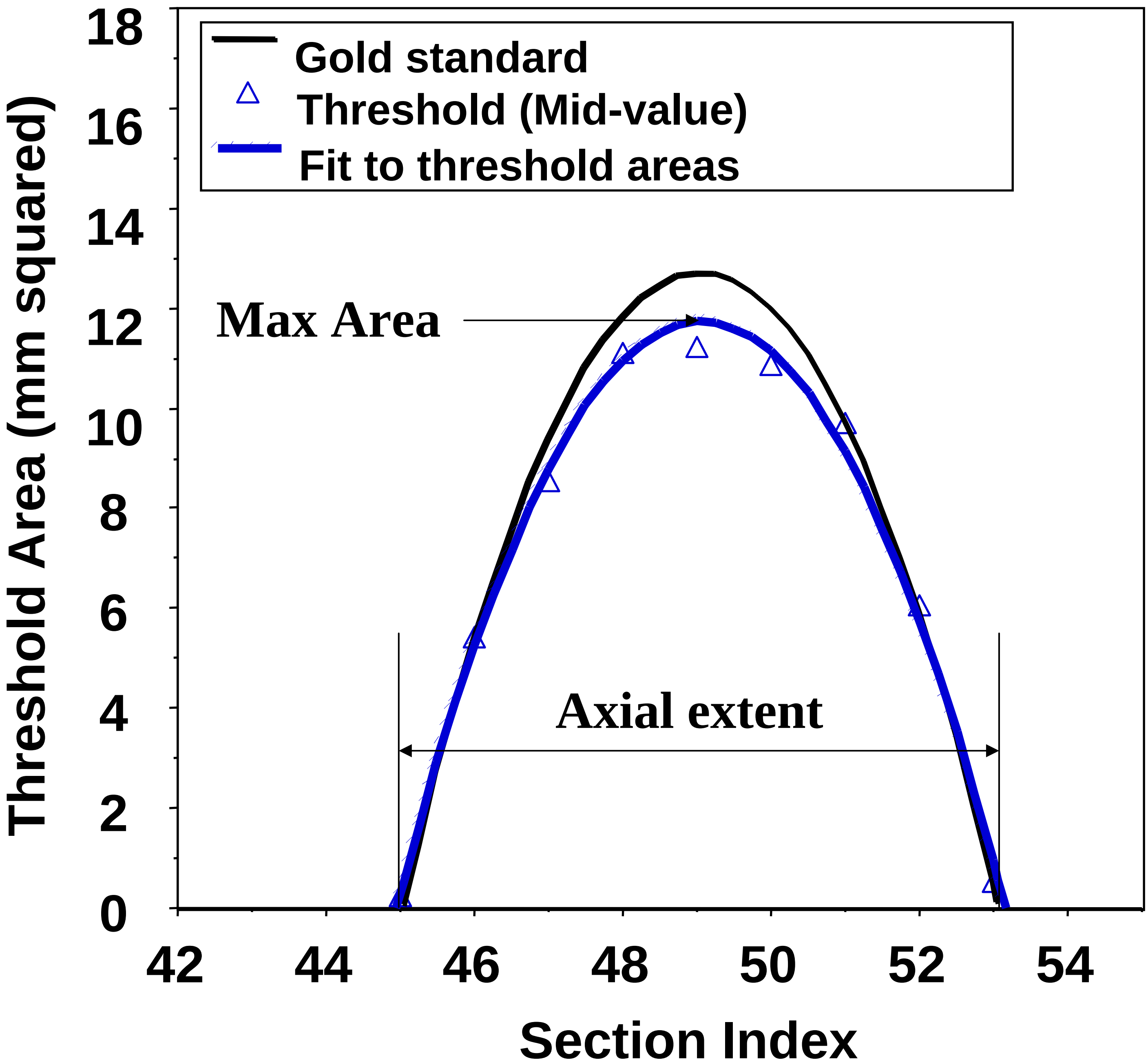


Figure 3 --Areas of a 4-mm nodule.

Comparison

- The volume of each feature was estimated on a common image-data set using 4 procedures.
- Diameter of each feature was estimated from nodule feature assuming that the feature was a sphere.

Relative bias

- Relative bias is just the “percentage” difference between the estimated and gold standard size. Relative bias is more significant than absolute bias, because absolute bias may be small just because the nodule is small. The important biological changes are relative changes. For example, a 1 mm change of 2 mm nodule is a 50% change, while it is only a 5% change of 20 mm nodule.

Four sizing procedure compared

Procedure 1: The threshold procedure using mid-value CT number

Procedure 2: From a surface representation of a 3-D rendering procedure using threshold segmentation (Able Software, 3D-Doctor, Lexington, MA): Threshold was set the same as Procedure 1.

Procedure 3: A 2-D multi-criteria thresholding procedure (Zhao-1999).

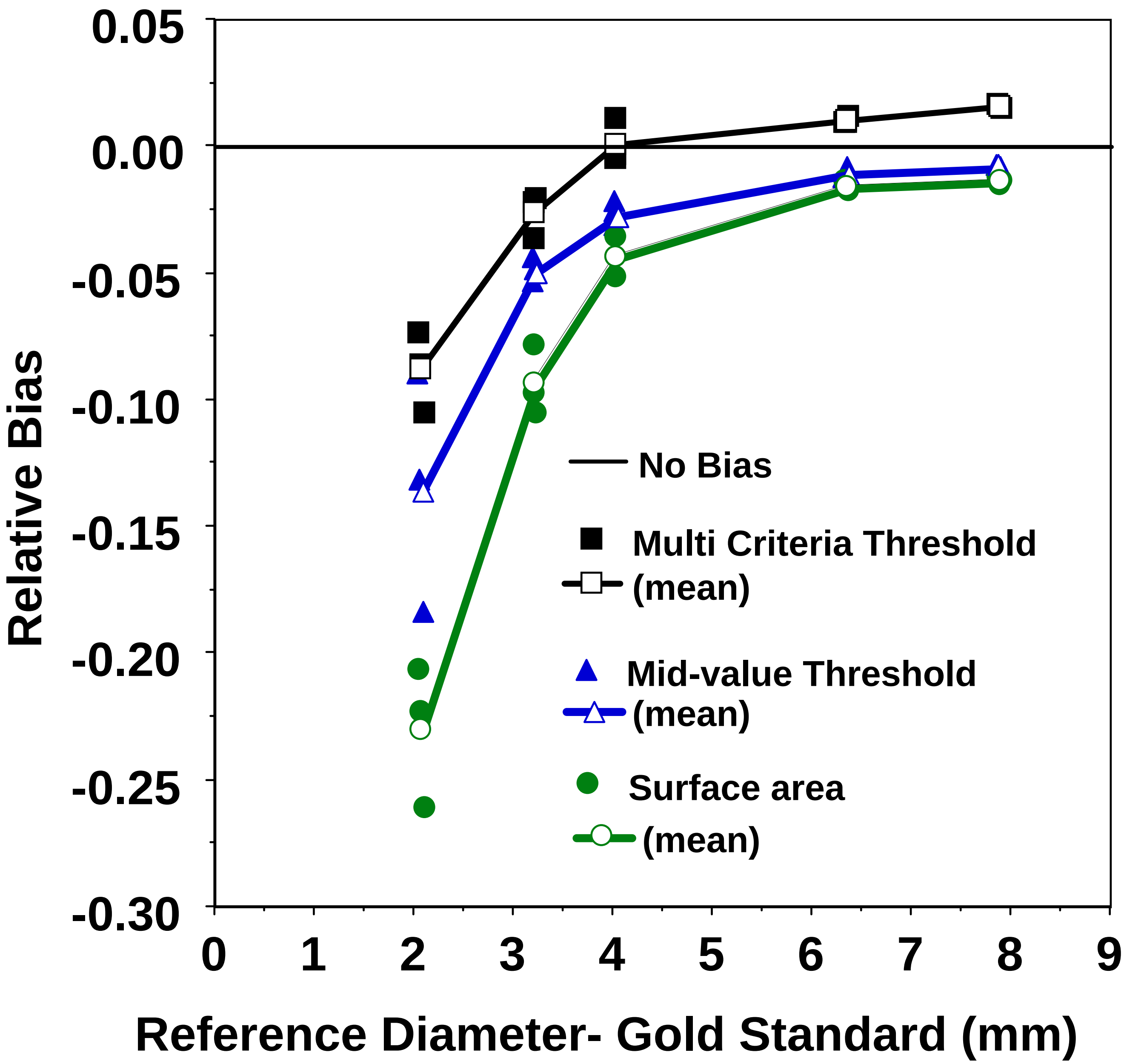
Procedure 4: Threshold value set by evaluation of CT number histograms (Kohl-2000): Its performance were the same as the threshold procedure, so results of this evaluation not presented.

Zhao B, Reeves A, Yankelevitz D, Henschke C. Two-dimensional multi-criterion segmentation of pulmonary nodules on helical CT images. Med Phys 1999; 26:889-895.

Kohl-2000, Unpublished results, Gerhard Kohl, Siemens Medical Systems, Forchheim, Germany

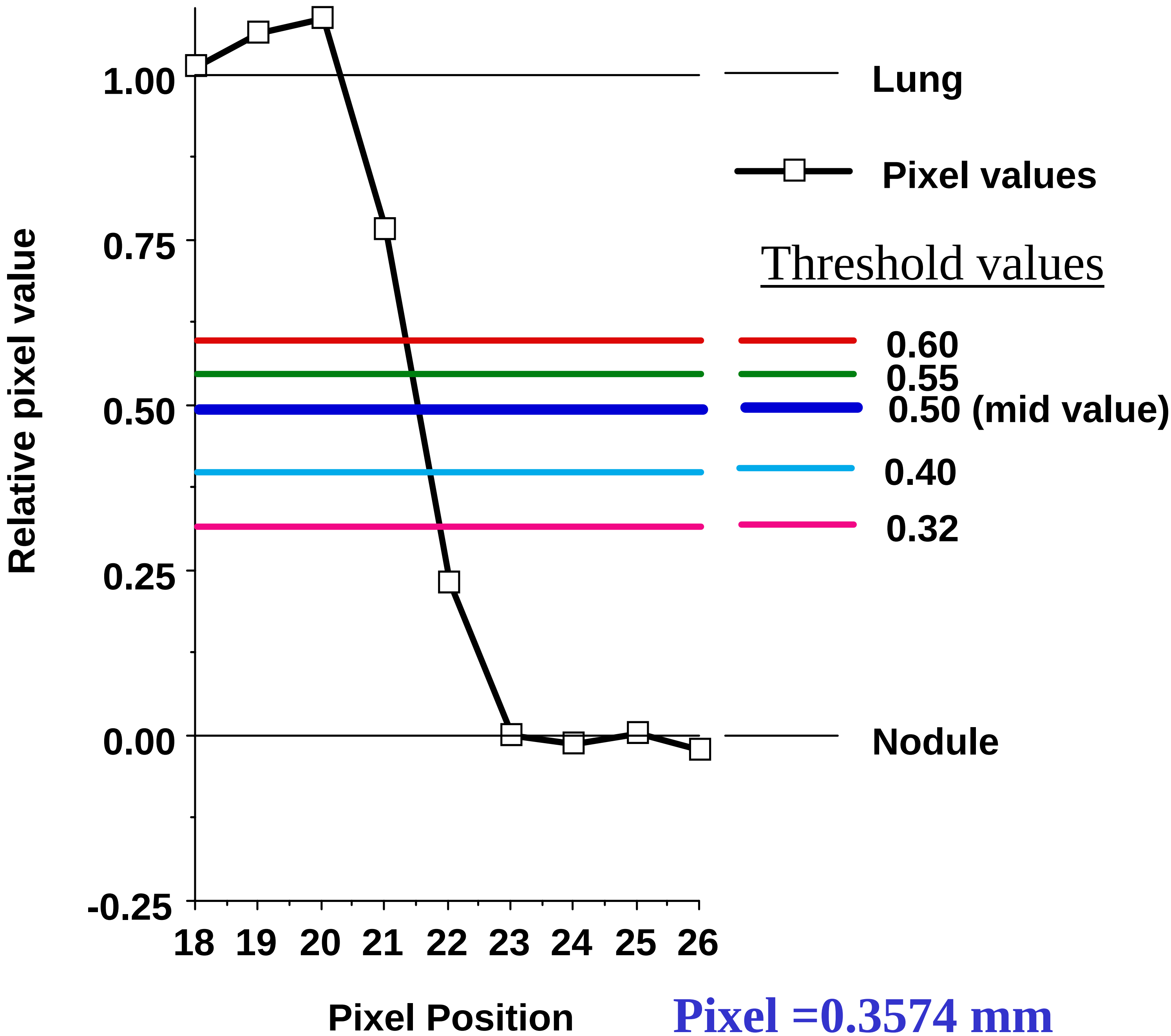
Figure 4

Relative Bias of 3 Sizing Procedures



Relative bias was estimated for each ball separately. The mean relative bias of the 3 balls of each size is plotted and connected.

Figure 5
Various thresholds plotted on edge



Pixel =0.3574 mm

Figure 6

Bias of Diameter Estimate from Volume Estimate (no fit)

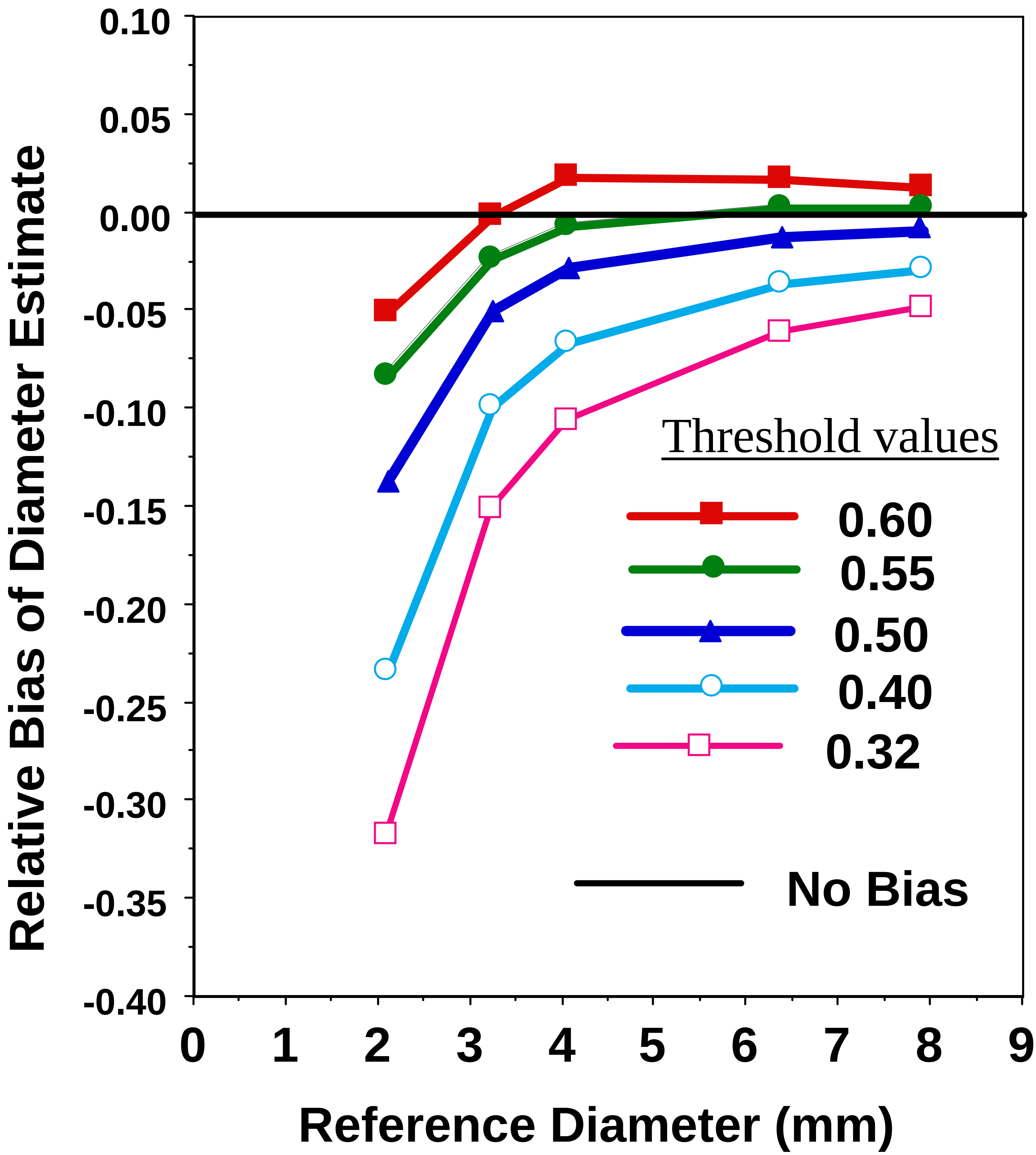


Figure 7

Bias of Diameter Estimate from Axial Extent (fit)

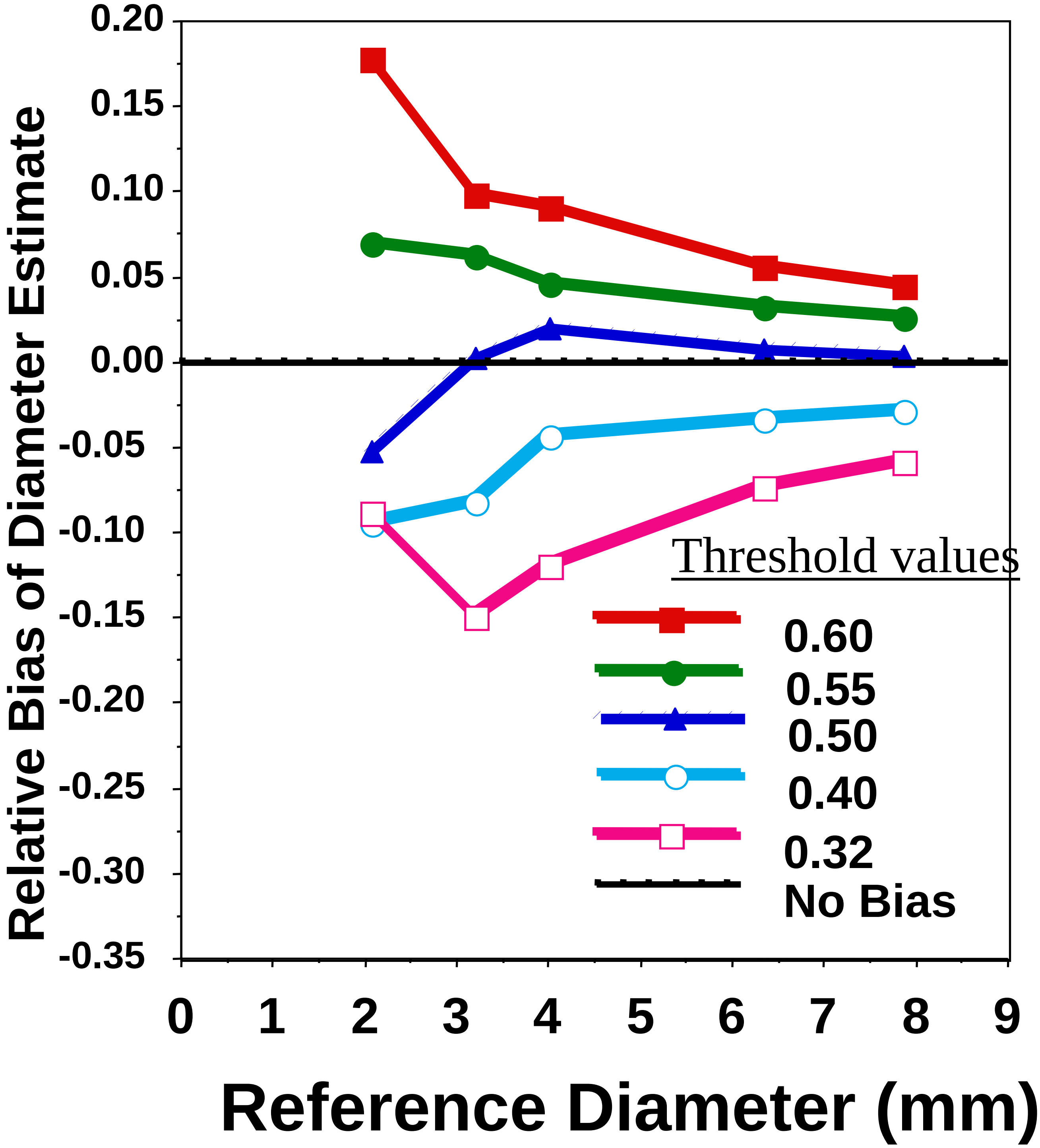
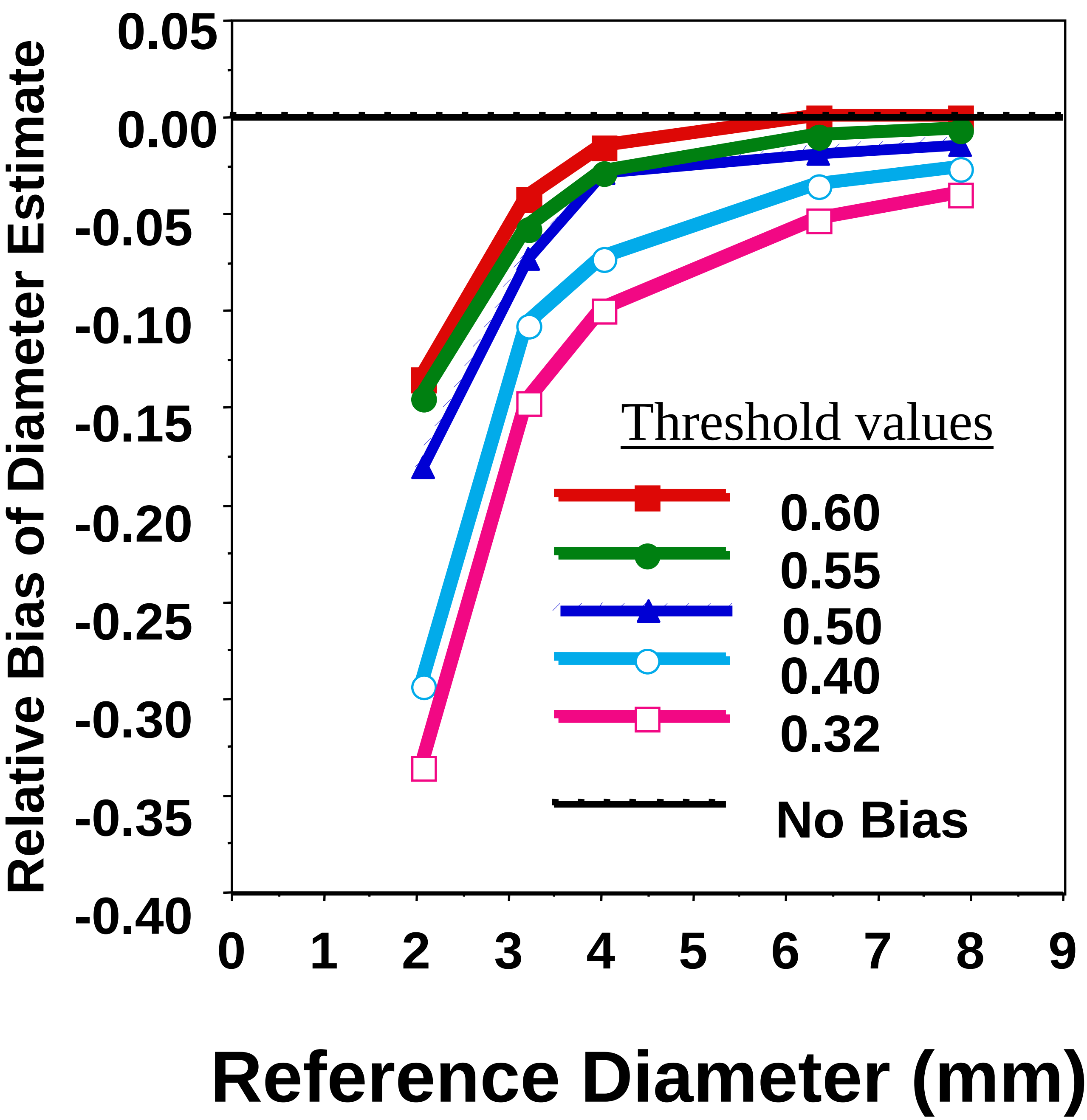


Figure 8

Bias of Diameter Estimate from Maximum Area (fit)



Results of Comparisons

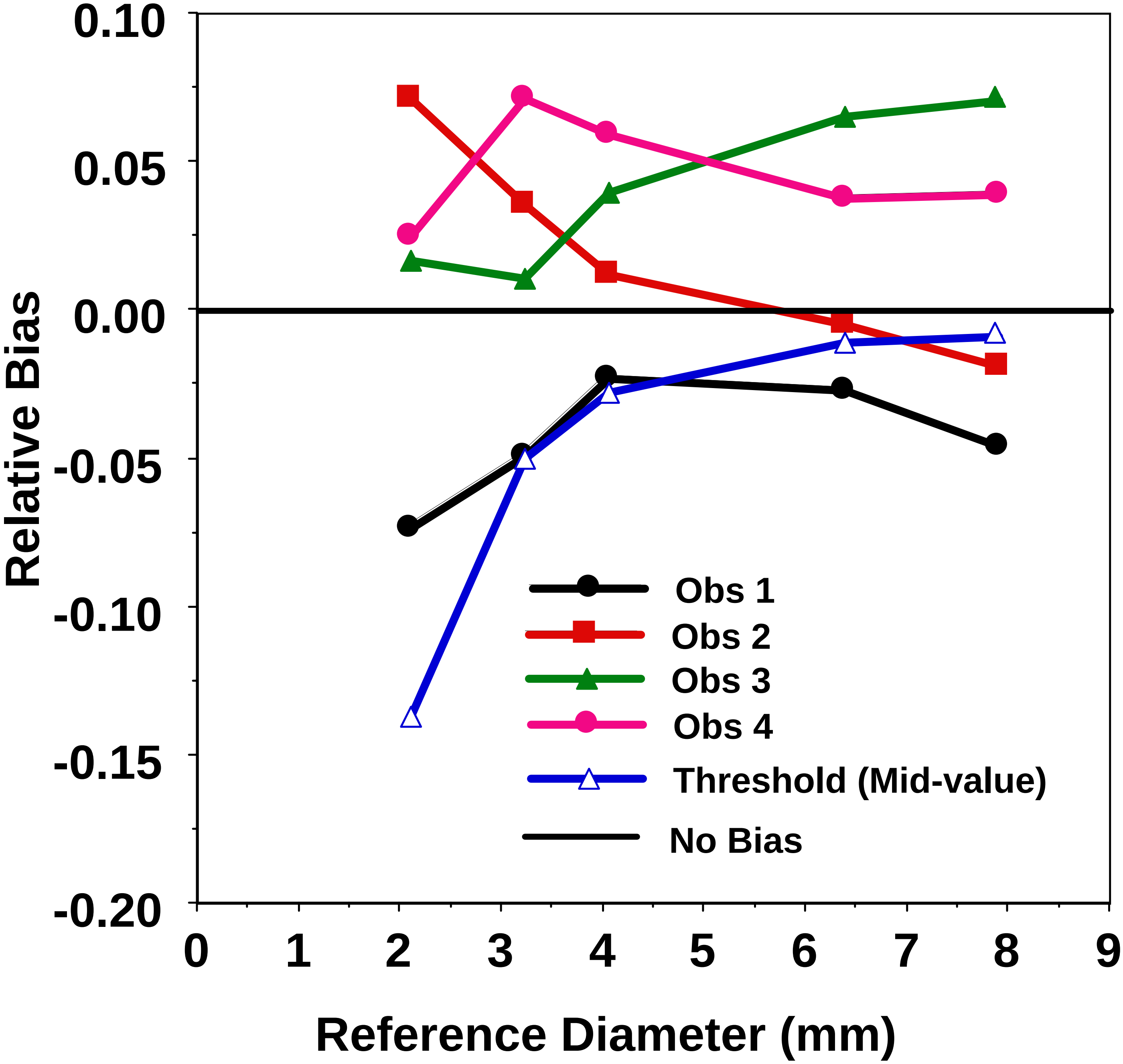
- Figure 4 shows that all procedures under-estimate the size of the smaller acrylic balls.
- Figure 4 shows the Mid-value Threshold under-estimates the size of all acrylic balls, while the Multi-Criteria Threshold procedure slightly over-estimates (2%) the size of larger acrylic balls.
- The small range of bias demonstrated in Figure 4 (less that 5%) for balls greater than 3.2 mm suggest that computer procedures applied to CT images of have potential for very precise measurement of nodule. Such procedures have the potential for detecting small changes of tumor volume.

Results of Comparisons

- The similarity of Figures 6 and 8, relative bias of the estimate of diameters from volume measurements and maximum area measurement indicate that the error in these procedures have the same source.
- Figure 6 suggests the reason that the Multi-Criteria Threshold overestimates that it overcompensates for effects of curvature and unsharpness by choosing a higher threshold.

Figure 9

Human Observer Measurement Same Images



Human Observer Measurement Same Images

- Figure 9 (from Jacobson-2000) are diameter estimates of 4 human observers who measured vertical and horizontal diameters on all sections on which the nodules were visible. A volumes were calculated and the diameters were estimated from the volume. This figure demonstrates effect of the edge criteria variation of humans.
- The image reference standard results (Figs. 6, 7, and 8) and additional human measurements (Fig, 9) suggest that human observers use a threshold criterion closer to lung CT numbers for image features in sections at the axial ends of the spheres.

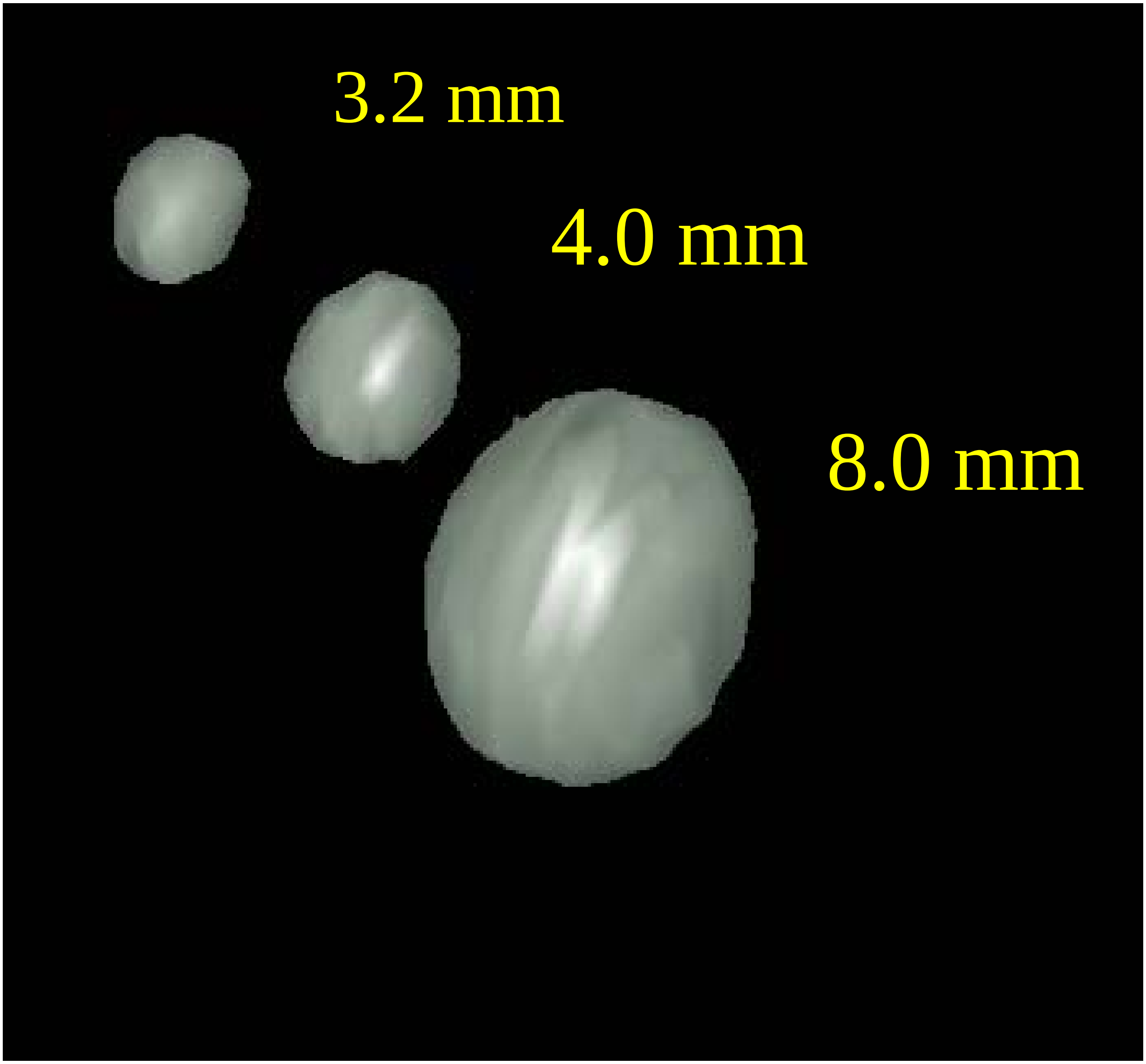
Jacobson-2000, this RSNA, Paper 187, page 243,   Validity of Objective and Subjective Measurements of Small Pulmonary Nodules from CT images

General Discussion

Response to comment: It's too simple.

- This phantom is just a first step.
- If a procedure doesn't work in a simple situation, why would one expect the procedure to work in complex medical background.
- It is useful for quality assurance protocols.
- Provides bound on performance for more complex backgrounds.
- Phantom is simple, consequently one can predict the biases and precision.
- The prediction can be used to optimize size estimation procedures and the optimizations verified.
- Evaluation of the results from reference standard provides insight into error mechanisms.

Image Surface Rendering



Note: Smoothness of surface renderings. The smoothing process decreased the nodule size that was estimated from this rendering.