Special Issue on

Cancer Screening with Computed Tomography (CT) Scan

Edited by:

Tao Li, MD, Ph.D.
University of Washington School of Medicine, USA

Review Article

Volumetric Analysis of Non-Calcified Lung Nodules with Thoracic CT: an Updated Review of Related work over the Last 5 Years

Marios A Gavrielides*, Qin Li, Rongping Zeng, Kyle J Myers, Berkman Sahiner, and Nicholas Petrick

Division of Imaging and Applied Mathematics, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, Maryland, United States

Abstract

Lung nodule volumetry with computed tomography has the potential for more reliable measurements of tumor size and therefore determination of temporal changes in a shorter interval of time. In a 2009 review, Gavrielides et al summarized the findings of studies that had examined the inter-related factors affecting the accuracy and precision of volumetric measurements of lung nodules with CT. In this review we update this earlier work by summarizing the recent body of literature in this field. In addition, we provide a list of publicly available resources for researchers and summarize current efforts towards building consensus and standardizing the use of volumetric CT.

INTRODUCTION

Improvements in Computed Tomography (CT) technology over the last decade have led to the development of three-dimensional (3D) methods for lung nodule volumetry, aiming to produce more accurate and consistent size measurements and therefore earlier determination of temporal changes. A number of inter-related factors can affect the accuracy and precision of volumetric CT, as discussed in a 2009 review by Gavrielides et al. [1]. These factors include scan acquisition and reconstruction parameters, lung nodule characteristics, tools used for volume estimation, and the interactions of such factors. The review

---

summarized the findings of a number of studies based on clinical as well as phantom and simulated data that had examined these factors. It was stressed that such effects needed to be understood and accounted for before volume could be fully utilized as a metric for estimating tumor burden in clinical practice. The review also stressed the need for public databases as resources for facilitating the assessment and comparison of different lung nodule size estimation methodologies with regards to measurement error.

Since the review was published, a number of new studies have added to the body of work examining factors affecting volumetric precision and accuracy. In this manuscript, we review the most significant findings of those studies that have appeared in the peer-reviewed literature. We also describe some publicly available resources for researchers in this field, summarize current efforts to build consensus and standardize the use of volumetric CT, and discuss some under-examined areas of research. Studies comparing volume with other measures of size, such as 1D and 2D measurements, are beyond the scope of this review.

Factors affecting uncertainty in volumetric measurements of lung nodules

CT acquisition and reconstruction parameters: Starting with acquisition parameters, *slice collimation* was identified as having a significant effect on volume estimation error in the 2009 review. Its significance was supported by findings in a phantom study by Gavrielides et al [2] of nodules with multiple shapes, density of 100HU, and size ranging from 5mm to 10mm in nominal diameter. Minimum detectable growth of nodule, based on estimated nodule volume, at a performance value of AUC=0.95 (AUC=Area under the receiver operating characteristic curve) for nodules with a baseline nominal diameter of 5mm was determined as 45% when using 16x1.5mm slice collimation, but improved to 17% when a thinner slice collimation protocol (16x0.75mm) was used. The effect for larger nodules was smaller, with a reduction in the order of 1-2%.

The effect of the product of tube current and exposure time (mAs) was not established consistently in studies discussed in the 2009 review. Soon after the review, Linning et al [3] examined the effect of various tube currents (ranging from 30 to 210 mA) on the accuracy of volumetric measurements of 13 artificial Ground-Glass Opacity (GGO) nodules (12 mm in diameter, cylindrical). Results derived from commercial software showed substantial underestimation of nodule volumes at ≤ 90mA and substantial overestimation of volumes at ≥120mA, but no statistically significant difference in absolute percentage errors. In a clinical study by Hein et al [4], inter-observer and intra-scan variability of volume measurements of 202 pulmonary nodules using CT data was compared at standard-dose (SD, 75 mAs) and ultra-low dose (ULD, 5mAs) using semi-automated segmentation software. Results showed mean relative differences of -0.7% and -0.2% between two readers for SD-CT and ULD-CT respectively. The authors concluded that the inter-observer variability of semi-automated volumetry of pulmonary nodules was independent of the dose level, indicating the feasibility of ultra-low dose scan CT. Similarly, no statistical significance was found when comparing

the limits of agreement in the intra-scan analysis. Dose was also not found to be a significant factor in the study by Gavrielides et al [5] described below.

In the 2009 review, *reconstructed slice thickness* was identified as one of the most important imaging parameters in terms of its effect on volumetric error, with its effect being more pronounced for smaller nodules. This finding was supported in a phantom study by Nietert et al [6] of 29 lobular solid nodules (3.0 to 15.9 mm). Results showed that uncertainty in estimating volume doubling time was highly dependent on slice thickness and that scans with slice thicknesses >2.5mm were essentially inadequate for detecting changes in nodule diameter in the order of 1mm. A significant effect of slice thickness was also reported in the phantom study by Prionas et al [7] of 55 spherical nodules (1.6 to 25.4mm) with mean density of ~118 HU embedded in gelatin background of mean density 37.6HU. Chen et al [8] reported a dependence of measurement *precision* on slice thickness and the interaction between slice thickness and nodule size, in a phantom study of spherical nodules (5 and 10mm in diameter). The influence of slice thickness on accuracy was limited. Slice thickness was also determined to be a significant factor in terms of measurement variability in a clinical study of 118 lung and liver lesions and lymph nodes [9].

In addition to its interaction with nodule size, the interaction of reconstructed *slice thickness with reconstruction kernel (or filter)* was also shown to be a significant factor. In a clinical study of 200 patients by Wang et al [10], CT data were reconstructed using three different parameter settings: 1mm and 2 mm slice thickness with a soft kernel, and 2 mm with a sharp kernel. Results showed that low-dose CT reconstructed with 1 mm section thickness and a soft kernel provided the most repeatable volume measurement in their study, and supported the use of consistent reconstruction settings for serial CT studies. A significant interaction of reconstruction kernel with slice thickness (as well as nodule size) was also reported by Prionas et al [7]. In related work on the effect of *reconstruction kernel*, Christe et al [11] reported that the volume measurement obtained with soft (B30) kernel was larger than that with hard (B70) kernel, with the magnitude of the difference dependent on the volumetric software (11.2% for one software and 1.6% for another software).

Reconstruction overlap was identified as an under-examined area of research in the 2009 review. Gavrielides et al [12] applied a matched filter-based volume estimation method in a phantom study of spherical synthetic nodules (5mm to 20mm in diameter, densities -630 and 100HU). Findings from this study showed that CT protocols that incorporated overlapping acquisition (slice increment equal to 50% of slice thickness) improve significantly both the precision and accuracy of volume estimates.

With the recent emergence of *iterative reconstruction algorithms* in clinical scanners due to the availability of larger computational capacities at the same cost and the ongoing efforts towards lower doses in CT [13], it is important to examine the effect of these algorithms on volume estimation. Chen et al [14] compared three reconstruction algorithms (Filtered-Back-Projection (FBP), and two iterative reconstruction algorithms) in terms of volume estimation for spherical nodules and varying dose levels (standard and reduced). Results from this study

showed statistically significant differences in accuracy between some combinations of reconstruction algorithm, slice thickness and software. In terms of Percent Bias (PB), these differences were relatively small (for slice thicknesses ≤1.25mm, differences in PB were in the order of ≤2% and ≤5% for 9.5mm and 4.8mm nodules respectively when using one clinical software for volume estimation, and ≤5% and ≤10% when using another clinical software). Precision was found to be generally comparable between FBP and the iterative reconstruction algorithms. Similarly, Willemink et al [15] reported comparable estimated volumes between FBP and iterative reconstruction for nodules ≥ 5 mm in diameter in a phantom study using a 256-detector CT scanner. The authors concluded that dose reductions up to 90.6% were possible without significant changes in nodule volumetry, even when using FBP.

The effect of *scanner type* was previously examined in a study of 16-detector row scanners [16]. Recently, Xie et al [17] assessed inter- and intra-scanner variability of pulmonary nodule volumetry acquired with two 64-detector row CT scanners and a common low-dose imaging protocol. In a phantom study of 5 spherical solid nodules (3mm to 12 mm, 100 HU) utilizing semi-automated software, no significant differences were found between the two CT scanners for nodules between 5-12 mm in diameter. Inter- and intra-scanner variability decreased with increasing nodule size.

Nodule characteristics: *Nodule size* was identified as one of the most important factors contributing to volumetric estimation error in the 2009 review, with multiple studies reporting increasing percent error with decreasing size. This result has been further supported in a number of new studies. Gavrielides et al [2] reported that minimum detectable growth, based on nodule volume estimates obtained using the same imaging protocol and software, was in the order of 40% for nodules with baseline nominal diameter of 5mm. This was reduced to approximately 15% as baseline nodule size increased to 9mm nodules. A similar pattern of decreasing error with increasing nodule size was also observed in a number of other studies [5,7,8,12,17].

An under-examined area of research reported in the 2009 review was the uncertainty in volumetric assessment of *sub-solid* or *non-solid* nodules. In the Linning study [3], the accuracy of volumetric measurements of GGOs was examined. Results showed overall Relative Percent Error (RPE) ranging from -22.7% to 17.3%. Significant effects included tube current and the radiodensity of nodule. Oda et al [18] also investigated volume estimation of GGO pulmonary nodules, through both phantom and clinical studies. Relative measurement error, calculated for the phantom study of spherical nodules (5mm to 12mm, densities -800 to -450 HU) ranged from -4% to 7%. In a separate clinical study of 59 nodules, the authors concluded that intra- and inter-observer agreement might be clinically acceptable for early detection of growth in GGO nodules that are 8mm in diameter or larger.

Nodule shape was one of the factors examined in the phantom study by Petrick et al [19] of 10 nodules (10, 20mm spherical, 20mm elliptical, 10mm lobulated, and 10mm spiculated nodules). Volumetric measurements were almost unbiased across all shapes with relative biases of radiologists using the 3D tool at

<2%. Precision was also good with estimated relative standard deviations of less than 10% for the same five nodule size/shape combinations.

Volume measurement method

Measurement method was shown to be an important contributor to volume estimation error in numerous studies summarized in the 2009 review. Different methods may interact differently with nodule characteristics such as vascular attachments, density and/or shape, image properties such as noise, and may require different levels of interaction with the operator (e.g. manual delineation vs. fully automated vs. semi-automated). In a clinical study by Ashraf et al [20] the reproducibility of three different segmentation algorithms within a semi-automated commercial volumetry software was assessed. Two readers at different sites read a set of 545 nodules from 488 CT scans collected at low-dose (40mAs) acquisition, slice thickness of 3mm, and a soft reconstruction kernel. Results showed differences between readers larger than 25% in only 4% of readings when the same segmentation algorithm option was used, and in 83% of readings (when different segmentation algorithms were used). The conclusion was that the same algorithm should be used to evaluate lung nodule volume for follow-up. In another clinical study, Christe et al [11] compared two different software packages (one automated and one semi-automated) for the volumetric measurement of 113 CT nodules from scans. The authors found a 42% mean volume difference between the two software for the same (B30) filter. The authors recommended the use of the same volumetric software with the same reconstruction filter for follow-up. In the phantom study by Chen et al [14] mentioned above, systematic differences were identified in the interaction of segmentation algorithms with different reconstruction algorithms. The authors emphasized the importance of extending current segmentation software to accommodate the image characteristics of iterative reconstructions [14].

In a study of inter- and intra-*observer variability* of CT tumor volume measurement in advanced Non-Small-Cell Lung Cancer (NSCLC) by Nishino et al [21], 53 lesions were measured by 2 radiologists using the same commercial software. Manual adjustment on automatically segmented boundary was allowed. One reader read all the cases once and another reader read all the cases five times. Results showed a relative mean difference of -3.7% between readers and -5.4% among measurements of the same reader, demonstrating high inter- and intra-observer agreement for the particular software application.

Other factors

Rampinelli et al [22] evaluated the reproducibility of automated volume measurements of 83 solid small nodules (5-10mm) in the same session but *separate breath-holds*. The authors concluded that volume variations greater than 30% for such nodules between two subsequent measurements should be confirmed by follow-up CT to confirm growth. In several other recent studies, parameters such as *field of view* [7], *kVp* [8] and *pitch* [8] were identified as items that did not significantly affect volumetric measurement error.

Summary of findings

The studies reviewed here identify a number of inter-related effects, including imaging parameters, nodule characteristics and measurement tools as having a significant impact on volumetric measurements. However, these findings were primarily obtained using different scanners, study designs, cases, measurement tools, procedures, and even performance metrics. These differences make it difficult to provide definitive conclusions regarding the requirements for the optimal way to use volumetric CT, as well as its limitations. However, it can generally be agreed from the increasing body of work in this field that: a) percentage measurement error increases with decreasing nodule size, b) slice thickness is a significant factor in the measurement of sub-centimeter nodules, c) the same measurement tool should be used to compare temporal scans, and d) temporal scans should preferably be acquired on the same scanner with the same slice collimation, and reconstructed with same slice thickness, slice overlap (increment) and kernel, when the goal of the scans is to provide quantitative lesion volume measurements.

In another review with relatively similar findings, Mozley et al [23] specifically concluded that *precision* in the measurement of clinical nodules was inversely proportional to slice thickness, directly proportional to the size of the mass, inversely proportional to the complexity of its shape, directly proportional to its contrast with surrounding tissue and dependent on factors related to selection and usage of measurement software.

DISCUSSION

There are several open issues related to the potential role of volumetric CT in clinical practice as a viable tool for nodule sizing. One issue is the standardization of imaging protocols for volumetric CT based on the findings reported in the literature, and forming consensus among stakeholders. The volumetric CT committee of the Quantitative Imaging Biomarkers Alliance (QIBA) was established towards this goal, aiming to test hypotheses about the technical feasibility and the medical value of imaging biomarkers, specifically through the example of volumetric imaging [24]. A number of collaborative projects were initiated through the consortium including a phantom study to compare volumetry with linear measurements of size [19] (QIBA-1A), a clinical reader study to compare reader performance (QIBA-1B), a multi-site phantom study to examine inter-scanner variability (QIBA-1C), and two public challenges to quantify inter-algorithm performance for both phantom and clinical data (QIBA-3A). A related QIBA committee has developed a framework for establishing the performance of quantitative imaging biomarker algorithms, including the treatise of study designs and statistical analysis approaches [25]. QIBA's efforts have resulted in the public release of *Profile* documents focusing on CT volumetry. Profile documents define a specific performance claim for an imaging biomarker and describe how the claim can be achieved through appropriate control of the image acquisition, patient, and measurement processes [24]. Accounting for measurement uncertainty will guide the effort to link volumetric assessment to clinical outcomes.

An important component toward qualification of volumetric CT for lung nodule sizing is the availability of shared datasets. A number of publicly available thoracic CT datasets that serve as valuable resources for researchers and developers of nodule volumetry tools are described in a review by Buckler et al [26]. Similarly, the availability of shared software for 3D data analysis is necessary; a number of open source tools are currently available including a lesion sizing toolkit reported by Krishnan et al [27], and the QI_Bench project (http://www.qi-bench.org) which provides a set of statistical tools for the assessment of quantitative medical imaging tools.

Other important issues to be determined include clinical workflow and the efficiency of volumetric tools. Tools will need to require minimal interaction with the observer and produce measurements quickly to avoid disturbing work flow. Finally, an under-examined area of research is the development of metrics that take into account changes in nodule *density* in addition to changes in volume since volume might not reflect treatment response such as a necrotic center. In related work, Sone et al [28] examined whether an increase in size and density of the Central Denser Zone (CDZ) in a nonsolid lesion could indicate progression or aggressiveness of lung cancer. In a preliminary study of three non-consecutive patients, they found excellent intra- and inter-reader agreement on volume and weight measurements of individual tissue portions stratified by density. In a follow-up analysis on a set of 85 patients, Sone et al [29] demonstrated that quantification of CDZ may be helpful in identifying nodules with problematic postsurgical clinical outcome.

In summary, we have provided an overview of recent study findings on the various factors affecting uncertainty in the assessment of lung nodule size with volumetric CT. Reducing measurement uncertainty can lead to earlier detection of change in nodule size, more efficient management of patient treatment, and a decrease in the size, cost, and length of clinical trials that examine the response of tumors to treatment.

REFERENCES

1. Gavrielides MA, Kinnard LM, Myers KJ, Petrick N. Noncalcified lung nodules: volumetric assessment with thoracic CT. Radiology. 2009; 251: 26-37.

2. Gavrielides MA, Li Q, Zeng R, Myers KJ, Sahiner B, Petrick N. Minimum detectable change in lung nodule volume in a phantom CT study. Acad Radiol. 2013; 20: 1364-1370.

3. Linning E, Daqing M. Volumetric measurement pulmonary ground-glass opacity nodules with multi-detector CT: effect of various tube current on measurement accuracy—a chest CT phantom study. Academic radiology. 2009; 16: 934-939.

4. Hein PA, Romano VC, Rogalla P, Klessen C, Lembcke A, Bornemann L, et al. Variability of semiautomated lung nodule volumetry on ultralow-dose CT: comparison with nodule volumetry on standard-dose CT. J Digit Imaging. 2010; 23: 8-17.

5. Gavrielides MA, Zeng R, Myers KJ, Sahiner B, Petrick N. Benefit of overlapping reconstruction for improving the quantitative assessment of CT lung nodule volume. Acad Radiol. 2013; 20: 173-180.

6. Nietert PJ, Ravenel JG, Leue WM, Miller JV, Taylor KK, Garrett-Mayer ES, et al. Imprecision in automated volume measurements of pulmonary nodules and its effect on the level of uncertainty in volume doubling time estimation. Chest. 2009; 135: 1580-1587.

7. Prionas ND, Ray S, Boone JM. Volume assessment accuracy in computed tomography: a phantom study. J Appl Clin Med Phys. 2010;

SciMedCentral

11: 3037.

8. Chen B, Barnhart H, Richard S, Colsher J, Amurao M, Samei E. Quantitative CT: technique dependence of volume estimation on pulmonary nodules. Phys Med Biol. 2012; 57: 1335-1348.

9. Zhao B, Tan Y, Bell DJ, Marley SE, Guo P, Mann H, et al. Exploring intra- and inter-reader variability in uni-dimensional, bi-dimensional, and volumetric measurements of solid tumors on CT scans reconstructed at different slice intervals. European journal of radiology. 2013; 82: 959-968.

10. Wang Y, de Bock GH, van Klaveren RJ, van Ooyen P, Tukker W, Zhao Y, et al. Volumetric measurement of pulmonary nodules at low-dose chest CT: effect of reconstruction setting on measurement variability. Eur Radiol. 2010; 20: 1180-1187.

11. Christe A, Bronnimann A, Vock P. Volumetric analysis of lung nodules in computed tomography (CT): comparison of two different segmentation algorithm softwares and two different reconstruction filters on automated volume calculation. Acta Radiologica. 2014; 55: 54-61.

12. Gavrielides MA, Zeng R, Kinnard LM, Myers KJ, Petrick N. Information-theoretic approach for analyzing bias and variance in lung nodule size estimation with CT: a phantom study. IEEE Trans Med Imaging. 2010; 29: 1795-1807.

13. Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in X-ray CT. Phys Med. 2012; 28: 94-108.

14. Chen B, Barnhart H, Richard S, Robins M, Colsher J, Samei E. Volumetric quantification of lung nodules in CT with iterative reconstruction (ASiR and MBIR). Med Phys. 2013; 40: 111902.

15. Willemink MJ, Leiner T, Budde RP, de Kort FP, Vliegenthart R, van Ooijen PM, et al. Systematic error in lung nodule volumetry: effect of iterative reconstruction versus filtered back projection at different CT parameters. AJR Am J Roentgenol. 2012; 199: 1241-1246.

16. Das M, Ley-Zaporozhan J, Gietema HA, Czech A, Mühlenbruch G, Mahnken AH, et al. Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. Eur Radiol. 2007; 17: 1979-1984.

17. Xie X, Willemink MJ, Zhao Y, de Jong PA, van Ooijen PM, Oudkerk M, et al. Inter- and intrascanner variability of pulmonary nodule volumetry on low-dose 64-row CT: an anthropomorphic phantom study. Br J Radiol. 2013; 86: 20130160.

18. Oda S, Awai K, Murao K, Ozawa A, Yanaga Y, Kawanaka K, et al. Computer-aided volumetry of pulmonary nodules exhibiting ground-glass opacity at MDCT. AJR Am J Roentgenol. 2010; 194: 398-406.

19. Petrick N, Kim HJ, Clunie D, Borradaile K, Ford R, Zeng R, et al. Comparison of 1D, 2D, and 3D nodule sizing methods by radiologists for spherical and complex nodules on thoracic CT phantom images. Acad Radiol. 2014; 21: 30-40.

20. Ashraf H, de Hoop B, Shaker SB, Dirksen A, Bach KS, Hansen H, et al. Lung nodule volumetry: segmentation algorithms within the same software package cannot be used interchangeably. Eur Radiol. 2010; 20: 1878-1885.

21. Nishino M, Guo M, Jackman DM, DiPiro PJ, Yap JT, Ho TK, et al. CT tumor volume measurement in advanced non-small-cell lung cancer: Performance characteristics of an emerging clinical tool. Acad Radiol. 2011; 18: 54-62.

22. Rampinelli C, De Fiori E, Raimondi S, Veronesi G, Bellomi M. In vivo repeatability of automated volume calculations of small pulmonary nodules with CT. AJR Am J Roentgenol. 2009; 192: 1657-1661.

23. Mozley PD, Schwartz LH, Bendtsen C, Zhao B, Petrick N, Buckler AJ. Change in lung tumor volume as a biomarker of treatment response: a critical review of the evidence. Ann Oncol. 2010; 21: 1751-1755.

24. Buckler AJ, Mozley PD, Schwartz L, Petrick N, McNitt-Gray M, Fenimore C, et al. Volumetric CT in lung cancer: an example for the qualification of imaging as a biomarker. Acad Radiol. 2010; 17: 107-115.

25. Obuchowski NA, Reeves AP, Huang EP, Wang XF, Buckler AJ, Kim HJ, et al. Quantitative Imaging Biomarkers: A Review of Statistical Methods for Computer Algorithm Comparisons. Statistical Methods in Medical Research. 2014.

26. Buckler AJ, Schwartz LH, Petrick N, McNitt-Gray M, Zhao B, Fenimore C, et al. Data sets for the qualification of volumetric CT as a quantitative imaging biomarker in lung cancer. Opt Express. 2010; 18: 15267-15282.

27. Krishnan K, Ibanez L, Turner WD, Jomier J, Avila RS. An open-source toolkit for the volumetric measurement of CT lung lesions. Opt Express. 2010; 18: 15256-15266.

28. Sone S, Tsushima K, Yoshida K, Hamanaka K, Hanaoka T, Kondo R. Pulmonary nodules: Preliminary experience with semiautomated volumetric evaluation by CT stratum. Acad Radiol. 2010; 17: 900-911.

29. Sone S, Hanaoka T, Ogata H, Takayama F, Watanabe T, Haniuda M, et al. Small peripheral lung carcinomas with five-year post-surgical follow-up: assessment by semi-automated volumetric measurement of tumour size, CT value and growth rate on TSCT. European Radiology. 2012; 22: 104-119.