Measurement of Focal Ground-glass Opacity Diameters on CT Images:

Interobserver Agreement in Regard to Identifying Increases in the Size of Ground-Glass Opacities

Ryutaro Kakinuma, MD, Kazuto Ashizawa, MD, Keiko Kuriyama, MD, Aya Fukushima, MD, Hiroyuki Ishikawa, MD, Hisashi Kamiya, MD, Naoya Koizumi, MD, Yuichiro Maruyama, MD, Kazunori Minami, MD, Norihisa Nitta, MD, Seitaro Oda, MD, Yasuji Oshiro, MD, Masahiko Kusumoto, MD, Sadayuki Murayama, MD, Kiyoshi Murata, MD, Yukio Muramatsu, MD, Noriyuki Moriyama, MD

Purpose: To evaluate interobserver agreement in regard to measurements of focal ground-glass opacities (GGO) diameters on computed tomography (CT) images to identify increases in the size of GGOs.

Materials and Methods: Approval by the institutional review board and informed consent by the patients were obtained. Ten GGOs (mean size, 10.4 mm; range, 6.5–15 mm), one each in 10 patients (mean age, 65.9 years; range, 58–78 years), were used to make the diameter measurements. Eleven radiologists independently measured the diameters of the GGOs on a total of 40 thin-section CT images (the first [n = 10], the second [n = 10], and the third [n = 10] follow-up CT examinations and remeasurement of the first [n = 10] follow-up CT examinations) without comparing time-lapse CT images. Interobserver agreement was assessed by means of Bland-Altman plots.

Results: The smallest range of the 95% limits of interobserver agreement between the members of the 55 pairs of the 11 radiologists in regard to maximal diameter was −1.14 to 1.72 mm, and the largest range was −7.7 to 1.7 mm. The mean value of the lower limit of the 95% limits of agreement was -3.1 ± 1.4 mm, and the mean value of their upper limit was 2.5 ± 1.1 mm.

Conclusion: When measurements are made by any two radiologists, an increase in the length of the maximal diameter of more than 1.72 mm would be necessary in order to be able to state that the maximal diameter of a particular GGO had actually increased.

Key Words: Lung; ground-glass opacity; computed tomography; diameter; interobserver agreement.

Since the advent of multislice computed tomography (CT) technologies and the widespread use of CT scanning of the chest for CT lung cancer screening as well as for clinical purposes, increasing numbers of focal ground-glass opacities (GGOs) are being detected (1–10). GGOs are defined as focal nodular areas of increased lung attenuation through which normal parenchymal structures such as airways, vessels, and interlobular septa can be seen (1,2). Most persistent GGOs are assumed to be pulmonary neoplasms (11,12). Although several computer-assisted volumetry studies of GGOs have been reported (13–16), measurement of GGO diameters on thin-section CT (TS-CT) images is a simple method of evaluating the size of GGOs in clinical practice. However, measurements of pulmonary nodules on TS-CT images have been found to vary with the radiologists (17). To our knowledge, no assessment of interobserver agreement in regard to diameter measurements of GGOs on TS-CT images and causes of its variability has ever been performed.

The purpose of this study was to evaluate interobserver agreement in regard to measurements of GGO diameters on TS-CT images as a means of identifying increases in the size of GGOs.

MATERIALS AND METHODS

The institutional review board approved this study, and informed consent was obtained from each patient.

Acad Radiol 2012; 19:389–394

From the National Cancer Center, Research Center for Cancer Prevention and Screening, 5-1-1 Tsukiji Chuo-ku, Tokyo, Japan (R.K., Y. Muramatsu, N.M.); Clinical Oncology Center, Nagasaki University Hospital (K.A.); Department of Radiology, Osaka Medical Center (K.K.); Graduate School of Biomedical Sciences, Division of Radiology and Radiation Biology, Nagasaki University (A.F., K. Minami); Department of Radiology, Niigata University Medical and Dental Hospital (H.I.); Department of Radiology, University of Ryukyu School of Medicine (H.K., S.M.); Department of Radiology, Niigata Cancer Center (N.K.); Department of Radiology, Komoro Kousei General Hospital (Y. Maruyama); Department of Radiology, Shiga University of Medical Science (N.N., K. Murata); Department of Diagnostic Radiology, Graduate School of Medical Sciences, Kumamoto University (S.O.); Department of Radiology, National Okinawa Hospital (Y.O.); Division of Diagnostic Radiology, National Cancer Center Hospital (M.K.). Received January 31, 2011; accepted December 1, 2011. **Address correspondence to:** R.K. e-mail: rkaki@ncc.go.jp

doi:10.1016/j.acra.2011.12.002

GGO Selection

Eleven GGOs, one each on TS-CT images of 11 patients, were randomly selected from the GGOs detected by the low-dose CT lung cancer screening project undertaken at our research center and followed at the outpatient clinic at the center. The GGO (size, 13.5 mm; mean CT value, −764 HU) on TS-CT image of a 63-year-old male was used for training in making diameter measurements by unblinded reading. The other 10 GGOs (mean size, 10.4 mm; range, 6.5–15 mm; mean CT value, −570 HU; range, −774 to 393 HU) in the other 10 patients (mean age, 65.9 years; range, 58–78 years) were used to make diameter measurements by blinded reading.

CT Scanning and Reconstruction Conditions

The CT scanner that was used for the follow-up CT examinations was an Aquilion 16 machine (Toshiba Medical Systems, Inc., Otawara, Japan). The follow-up CT examinations consist of scanning of the entire lungs and additional TS-CT of the GGOs. The entire lungs were scanned using following protocol: 120 kV, RealEC SD20, beam collimation 1 mm × 16, pitch factor 0.69 and 0.5 second/rotation, and the CT images were reconstructed at 5-mm intervals, a slice thickness of 5 mm, a field of view (FOV) of 32 cm, and a lung algorithm (FC10). TS-CT of the GGOs was performed with a scan range of 4 cm under the following protocol: 120 kV, 300 mA, beam collimation 0.5 mm × 16, pitch factor 0.69, and 0.5 second/rotation, and the TS-CT images were reconstructed at 1-mm intervals, a slice thickness of 1 mm, an FOV of 22 cm, and super high-resolution sharp lung algorithm (FC82).

Reading and Measurement Settings

The local area network used in this study consisted of two personal computers, two monitors for each computer (a 1.3-megapixel liquid crystal monochrome display monitor for data input and a 9-megapixel liquid crystal monochrome display monitor for viewing the GGOs on the TS-CT images and measuring), and a server.

There were 11 observers. All were thoracic radiologists, and their chest CT reading experience ranged from 6 to 27 years (mean, 17 years). The training GGO was presented to the observers three times in succession and measured three times in succession by each observer in an unblinded setting (ie, the observers were well aware that they were reading the same GGO three times). To allow familiarization with the reading system, they were asked to click the start point and end point of the maximal diameter to measure the maximal diameter of the GGO three times on the same TS-CT images. Then, a total of 40 TS-CT images—that is, 10 TS-CT images from the first follow-up CT examination (Group 1), 10 TS-CT images from the second follow-up CT examination (Group 2), 10 TS-CT images from the third follow-up CT examination (Group 3), and the same 10 TS-CT images as in Group 1 (Group 4)—were presented to each observer randomly in a blinded setting. The maximal diameter and perpendicular diameter of each GGO were measured at four-fold magnification of the TS-CT images. The X, Y, and Z coordinates of each GGO on the TS-CT images and their maximal diameter and perpendicular diameter were output to the database in the server. The X and Y coordinates of the start points and the end points of the maximal diameters and perpendicular diameters were also recorded in the server database.

Statistical Analysis

We evaluated interobserver and intraobserver agreement between the members of pairs (n = 55) of the 11 observers in regard to the GGO diameter measurements on the 40 TS-CT images (Groups 1, 2, 3, and 4). Interobserver and intraobserver agreement was assessed by means of Bland Altman plots (18) drawn with the MedCalc software program (MedCalc Software, Mariakerke, Belgium). Descriptive statistics regarding agreement between CT slice selection by the 11 radiologists (Groups 1, 2, 3, and 4), distances between click points as start points or as end points of the maximal diameters of GGOs made by members of any pair of the 11 radiologists (Groups 1, 2, 3, and 4), and differences in angle between the maximal diameters of each GGO measured by the members of any pair of the 11 radiologists (Groups 1, 2, 3, and 4) were calculated by using software programs (JMP, SAS Institute, Cary, NC; Excel, Microsoft, Redmond, WA). When assessing the distances between click points as start points or as end points of the maximal diameters and differences in angle between the maximal diameters, we assumed that the diameter measurements by the 11 radiologists were performed on the same slice of each GGO.

RESULTS

Interobserver and Intraobserver Agreement in Regard to Diameter Measurements

The results are shown in Table 1. The smallest range of the 95% limits of interobserver agreement between the members of the 55 pairs of the 11 radiologists in regard to maximal diameter in Groups 1, 2, 3, and 4 (Figs 1, 2) was −1.14 to 1.72 mm, and the largest range was −7.7 to 1.7 mm. The mean value of the lower limit of the 95% limits of agreement was −3.1 ± 1.4 (SD) mm, and the mean value of their upper limit was 2.5 ± 1.1 mm.

Interobserver and Intraobserver Percentage Agreement in Regard to Slice Selection

Interobserver agreement in regard to the slice on which the radiologists thought the area of the GGO appeared to be maximal ranged from 30% to 100% (mean ± SD, 58 ± 17%) in regard to the images in Group 1, from 30% to