ORIGINAL ARTICLE

Image Subtraction Facilitates Assessment of Volume and Density Change in Ground-Glass Opacities in Chest CT

Marius Staring, PhD,* Josien P. W. Pluim, PhD,* Bartjan de Hoop, MD,† Stefan Klein, PhD,*
Bram van Ginneken, PhD,* Hester Gietema, MD,† George Nossent, MD,‡ Cornelia Schaefer-Prokop, MD,§
Saskia van de Vorst, MSc,† and Mathias Prokop, PhD†

Objectives: To study the impact of image subtraction of registered images on the detection of change in pulmonary ground-glass nodules identified on chest CT.

Materials and Methods: A cohort of 33 individuals (25 men, 8 women; age range 51–75 years) with 37 focal ground-glass opacities (GGO) were recruited from a lung cancer screening trial. For every participant, 1 to 3 follow-up scans were available (total number of pairs, 84). Pairs of scans of the same nodule were registered nonrigidly and then subtracted to enhance differences in size and density. Four observers rated size and density change of the GGO between pairs of scans by visual comparison alone and with additional availability of a subtraction image and indicated their confidence. An independent experienced chest radiologist served as an arbiter having all reader data, clinical data, and follow-up examinations available. Nodule pairs for which the arbiter could not establish definite progression, regression, or stability were excluded from further evaluation. This left 59 and 58 pairs for evaluation of size and density change, respectively. Weighted kappa statistics (κ_w) were used to assess interobserver agreement and agreement with the arbiter. Statistical significance was tested with a κ z-test.

Results: When the subtraction image was available, the average interobserver κ_w improved from 0.52 to 0.66 for size change and from 0.47 to 0.57 for density change. Average agreement with the arbiter improved from 0.61 to 0.76 for size change and from 0.53 to 0.64 for density change. The effect was more pronounced when observer confidence without the subtraction image was low: agreement improved from 0.26 to 0.57 and from 0.19 to 0.47 in those cases.

Conclusions: Image subtraction improves the evaluation of subtle changes in pulmonary ground-glass opacities and decreases interobserver variability.

Key Words: chest CT, pulmonary nodules, evaluation of change

(*Invest Radiol* 2009;44: 61–66)

Differentiation of benign and malignant nodules in lung cancer screening trials[1–4] is largely based on the detection of subtle density and size changes after a short period of time. Various software tools are available for automatic or semiautomatic volumetry. Multiple studies have proven the superiority of those software tools over manual measurements of diameters[5] and visual assessment. Until now, these software tools are only approved for solid nodules and not for ground-glass or part-solid nodules.

The prognostic significance of these so-called ground-glass opacities (GGOs) is not yet fully understood. They may represent, for example, infections, scarring, fibrosis, atypical adenomatous hyperplasia, bronchoalveolar carcinoma (BAC), or minimally invasive adenocarcinoma.[6–8] The Early Lung Cancer Action Project (ELCAP) group reported a malignancy rate of 34% in nonsolid lesions and a 63% malignancy rate in part-solid nodules during their lung cancer screening trial.[9] The histology of malignant part-solid or nonsolid nodules was predominantly BAC or adenocarcinoma with bronchoalveolar features.[9] However, the differential diagnosis of such nonsolid nodules also includes a number of benign lesions that can either regress over time or remain stable.

The decision to treat a patient with a GGO depends on the level of suspicion for malignancy. Current ACCP guidelines[10] recommend that surgical biopsy is used to establish a histopathologic diagnosis in all patients with suspected bronchoalveolar carcinoma (BAC), a carcinoma known to frequently present as a GGO or part-solid nodule. The 2 most important criteria for increasing suspicion and further invasive work-up are change in size and the appearance of a solid component. In case of a BAC, the solid part visible on CT represents the collapse of the alveoli, subsequent formation of a fibrotic focus, and proliferation of tumor cells. Therefore, the presence of a solid part within the nodule on CT generally indicates invasive growth.[11] Consequently, the early detection of size and density change of a GGO may result in an early detection of malignancy and is commonly taken as an indication for surgical resection.[11–13] Nakamura et al reported that early detection and treatment of GGOs improves the prognosis of lung cancer.[14]

Detection of growth or density change of GGOs, however, is frequently difficult, especially when the lesion has a complex shape or when the position of the lesion varies between scans because of different levels of inspiration. Although techniques for automated volume measurements have been extensively studied for solid nodules,[15–18] such routines are just emerging for nonsolid GGOs.[19] Diameter measurements have been shown to suffer from low reproducibility in solid nodules[5] and therefore are even more difficult to use in GGOs. The current clinical practice for detecting change of GGOs is a side-by-side visual comparison of follow-up scans, aided by manual diameter measurements if appropriate.

In this manuscript, we suggest an alternative approach that provides a subtraction image along with the original images to more readily appreciate changes. To reduce artifacts in the subtraction image because of motion, and to get a meaningful subtraction image, the scans were locally matched before subtraction. The goal of this study was to investigate whether subtraction of such registered images improves detection of subtle changes in pulmonary ground-glass opacities. The ultimate goal is to identify malignant nodules by depicting growth in size and/or density.

MATERIALS AND METHODS

Study Participants

A cohort of 33 individuals (25 men, 8 women; age range 51–75 years; mean 59 years) with 37 focal ground-glass lesions (29 participants with 1 nodule, 4 with 2) were recruited from a multi-center lung cancer screening trial. The NELSON study[4,20] was approved by the Dutch Ministry of Health and by the institutional review board of each participating center. Written informed consent had been obtained from each participant. All participants smoked a minimum of 16 cigarettes/d for 25 years or 11 cigarettes/d for 30 years. A total of 2994 volunteers were scanned between April 2004 and February 2007 in one of the participating centers. From this group, the participants with a GGO or part solid nodule visible in at least 2 consecutive scans were included in our study. This meant that we could include 16 patients with 2 scans, 14 patients with 3 scans, and 3 patients with 4 scans. The time interval between scans was typically 3 or 9 months (mean: 6 months), with a minimum of 2 and a maximum of 13 months. A ground-glass opacity was defined as a focal area of increased lung opacity that did not distort or obscure the underlying lung markings (bronchi and lung vasculature). A part-solid nodule was defined as a nodule that contained components of soft-tissue attenuation in addition to ground-glass components. The nodule locations were determined as part of the lung cancer screening trial and were available for this study.

Mean initial nodule size as determined by manually measuring the maximum nodule diameter in sagittal, coronal, or transversal direction was 14 mm (standard deviation: 6.7 mm; range: 6–34 mm; 11 in the range 5–10 mm, 21 in 10–20 mm, 5 nodules >20 mm). The number of pure GGOs was 24; 13 nodules were part-solid. Histologic diagnosis was available for 2 of the included nodules and revealed adenocarcinoma for both GGOs.

CT Imaging

Imaging data were acquired with low-dose CT, using 16-detector-row spiral CT scanners (Mx8000 IDT or Brilliance 16P; Philips Medical Systems, Cleveland, OH). Data were acquired in approximately 12 seconds in helical mode with 16×0.75 mm collimation and 15-mm table feed per rotation (pitch 1.3). A caudo-cranial scan direction was chosen to minimize breathing artifacts. Participants were asked to take a deep breath and to hold their breath. No intravenous contrast material was injected. Exposure settings were 30 mAs at 120 kVp for patients weighing less than 80 kg (15/33), and 30 mAs at 140 kVp for those weighing more than 80 kg (18/33), without dose modulation. All participants received the same radiation dose for all follow-up scans. The in-plane voxel size ranged from 0.53×0.53 mm to 0.84×0.84 mm. Transverse images were reconstructed at a thickness of 1.0 mm and a 0.7-mm increment. All scans were reconstructed with a 512×512 matrix and a moderately soft filter (Philips B).

Registration

Image registration is the process of spatially aligning 2 images, so that corresponding structures overlap. The CT scans were registered using the mutual information measure,[21] using the software package elastix, freely available from http://elastix.isi.uu.nl. Details of the registration method are provided in Staring et al.[22] The follow-up scan was deformed to match the image acquired at an earlier date. For improved registration performance, only the lungs were registered. The lungs were automatically segmented using an algorithm described by Sluimer et al.[23] The result was used as a mask during the registration, so that only the lungs were matched. To prevent the nodules from deforming in a nonrigid fashion, which would effectively conceal nodule growth, a local rigidity penalty term was employed.[22] This method allows the lungs to be registered

nonrigidly, whereas the nodule is not compressed or enlarged. The region that has to be kept rigid needs to be defined, but only roughly. Therefore, a 3-dimensional ellipsoid was manually placed around the ground-glass opacity on the follow-up scan. The size and shape of the ellipsoid were changeable according to the individual nodule size and shape. Because the delineation of the GGO only needs to be rough,[22] it is expected that the placement of the ellipsoid has a negligible effect on the registration result. The subtraction image was constructed by subtracting the baseline image from the deformed follow-up image. This way, growth is observed as a white rim, shrinkage as a dark rim. Density increase is observed in the subtraction image as a bright area, density decrease as a dark area. The region outside the lungs was hidden on the subtraction image to focus observer attention on the nodule.

Observer Study

Pairs of scans of the same GGO were presented on a computer screen, with or without an additional subtraction image. Figure 1 shows an example of the data. To optimize the display, the 3D CT data were cropped to a region of $80 \times 80 \times 80$ mm around the nodule, ensuring that the observers did not have to locate the nodule. A subtraction image was created after registering the image pairs using the method described above.

Observers were asked to independently rate volume and density change of the ground-glass opacity and to rate the confidence in their decision. Volume changes were rated on a 3-point
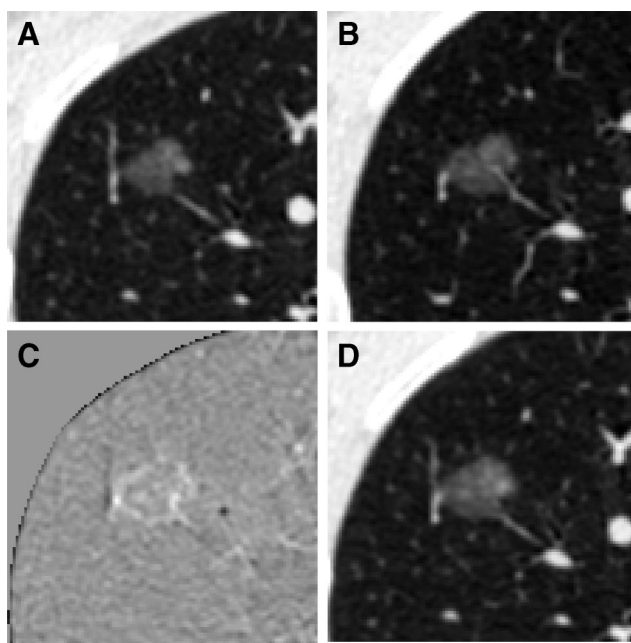
FIGURE 1. Example of a GGO that grew on an 8-month follow-up scan. The observers compared baseline (A) to follow-up (B). In one test set, no subtraction image was provided; in the other test set, the subtraction image (C) was available. For the standard of reference, a registered follow-up data set (D) was also available in addition to data sets (A–C). Note that the nodule is not displayed exactly identical on (A) and (B) because of small differences in inspirational depth. Also note the white rim on (C), indicative of tumor growth. For this particular lesion, all observers found an increase in size, both with and without the subtraction image. This had been confirmed by the arbiter with high confidence.
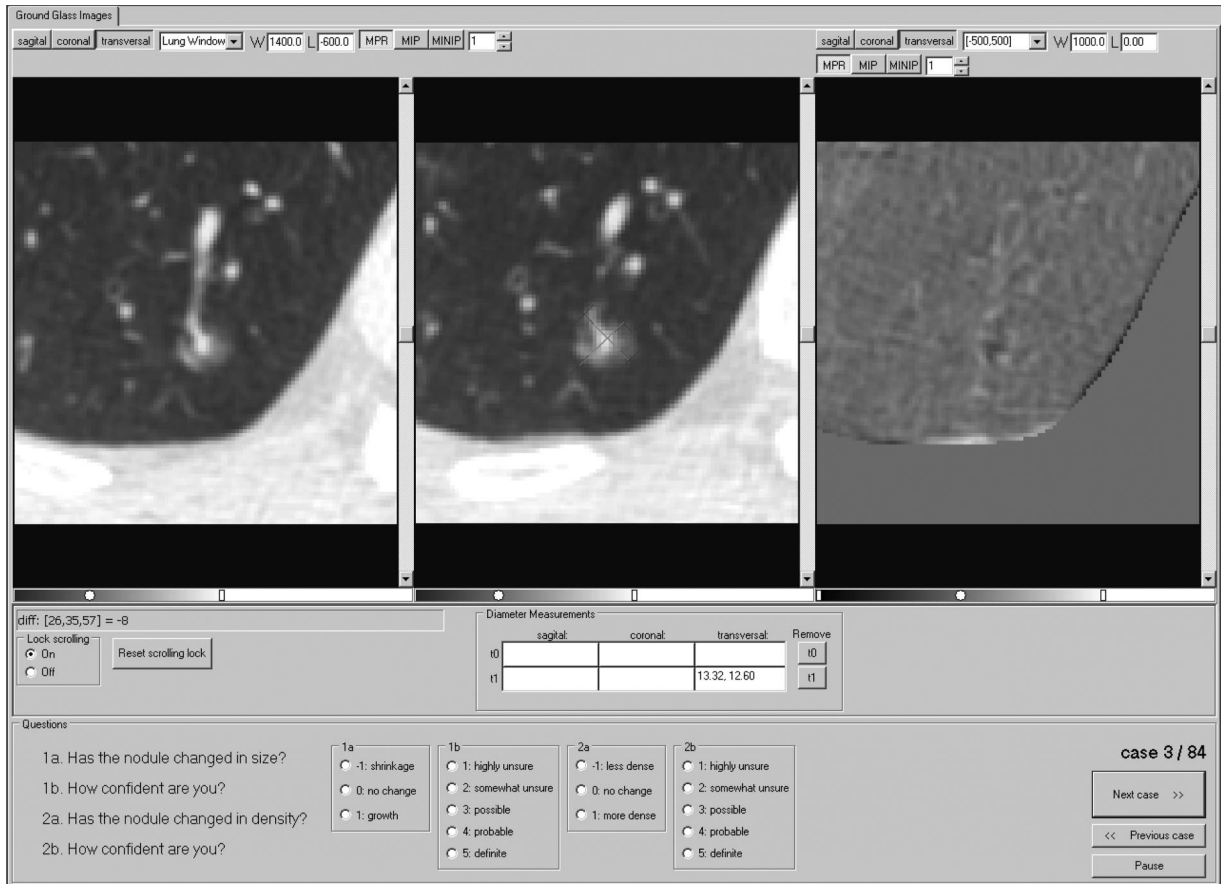
FIGURE 2. A screen shot of the interface. The middle image is a follow-up of the data set on the left. In this case, the observer is also provided with the subtraction image after registration.

scale as either shrinkage, no change, or growth. Density changes were also evaluated on a 3-point scale as either less dense, no change or more dense. Confidence was rated on a 5-point scale, from highly unconfident to highly confident. A user interface was developed to perform the observer study (Fig. 2). The user interface implemented standard viewing capabilities, such as zooming and adjustment of the window level settings. The default window level was a lung window (window: 1400 HU; level: −400 HU). As in a clinical situation, the observer could also make manual diameter measurements of the GGO, if warranted. Before the start of the study, a short (10–15 minute) training session was performed on solid nodule data to familiarize the observer with the software.

Per patient all combinations of image pairs of the same nodule, not counting duplicates, were shown to the observers. For instance, if 3 time points (t_0, t_1, t_2) were available, the pairs (t_0, t_1), (t_0, t_2) and (t_1, t_2) were evaluated. This resulted in 1 pair of scans for 18 nodules, 3 pairs of scans for 16 nodules, and 6 pairs of scans for 3 nodules, a total of 84 combinations. Each pair was shown once without the subtraction image and once with. Both scans were shown in their original format, ie, not registered to each other, as in current clinical practice.

The 84 image pairs were shown to observers in random order, determined by a computer program. If there was more than 1 image pair for a nodule, then the various pairs were treated as separate nodules. The observers were blinded to all patient information; it was unknown to the observers which pairs corresponded to the same nodule. All image pairs were scored twice by each observer, once

with and once without the subtraction image. Each observer scored in 2 reading sessions, which were at least 4 weeks apart. Per session, all image pairs were scored, half of them with subtraction image and half without. In the second session, all image pairs that had been evaluated without subtraction image in the first session were now presented with subtraction image and vice versa. Per image pair, the reading time was measured automatically, a fact the observer was informed of.

Four observers with substantial experience in chest CT or in reading lung cancer screening scans evaluated the image pairs: a radiology resident with 3 years of experience with CT lung cancer screening (>2000 scans), a pulmonologist with more than 15 years of experience with chest CT, a chest radiologist with more than 15 years of experience in CT, and a CT technician with special training in evaluating and reporting cancer screening CTs (>3000 scans in 2 years).

Determining a standard of reference poses the following challenges: a malignant lesion does not necessarily grow within a certain time interval. A benign lesion does not necessarily remain completely stable over time. Histopathology thus provides no true gold standard for the imaging task of determining size or density change. We therefore chose the following approach.

The standard of reference was established by an independent chest radiologist with more than 15 years of experience who did not participate in the observer study. He served as an arbiter who had access to the anonymous opinions of all observers. In addition, he was provided with all other available information including a fol-

low-up scan registered to the baseline scan. In case of multiple image pairs from the same nodule, he could view the pairs in the correct order. This way he could more easily follow disease progress. This arbiter was asked to state whether the confidence level for his final ruling was definite or not.

Despite availability of all this information, the reference observer was not able to confidently establish the presence of changes in all cases. To minimize impact of uncertainty on data analysis, only cases for which the standard of reference could be established with the highest confidence (definite) were included for further data analysis. Consequently, 59 of 84 nodule pairs were eligible for evaluation of size change and 58 pairs were eligible for evaluation of density change. This final group consisted of 30 nodules (20 GGO and 10 part-solid).

Statistical Analysis

Agreement with the standard of reference and between observers was analyzed with the weighted kappa statistic κ_w.[24,25] Weights are taken inversely proportional to the difference in ratings.[24–26] To statistically compare κ_w with a value $\hat{\kappa}$, the κ z-test was employed. Note, that this test is not paired and that it is therefore conservative. A paired test, if it were available, could indicate statistical significance sooner.

Besides measuring the overall effect of image subtraction, we also investigated the effect for difficult and least difficult cases separately, as indicated by observer confidence. Cases for which an observer was highly confident (marked 5) were subgroup 1, cases marked as 4 or lower were subgroup 2. For each observer, the agreement with the standard of reference was calculated for both subgroups separately.

Differences in confidence between the 2 methods (without and with subtraction) were assessed by a 2-sided paired Wilcoxon test.

A value of $P < 0.05$ was considered to indicate a statistically significant difference. Statistical analyses were performed with Microsoft Office Excel 2003 software (Microsoft, Redmond, WA).

RESULTS

According to the reference standard, 5/30 (17%) GGOs showed regression over time and no change was detected in 10/30 (33%) GGOs. Fifteen nodules (50%) showed an increase in size, whereas 8 of these GGOs (53%) additionally showed an increase in density. The observers needed on average 75 seconds to score a case, both without and with subtraction. Two of the observers needed less time with the subtraction image (from 75 seconds to 70 seconds and 57 seconds $\rightarrow$ 33 seconds) and 2 needed more time (71 seconds $\rightarrow$ 84 seconds and 96 seconds $\rightarrow$ 113 seconds).

The 4 observers made 14, 48, 4, and 16 (on average 21) diameter measurements, in all 84 image pairs when no subtraction image was provided. With a subtraction image, these numbers changed to 0, 0, 6, and 10 (on average 4).

The average observer confidence ratings for growth detection were 4.3 (of 5) for the standard technique and 4.4 with the subtraction image available (significant increase for 2 observers: $P = 0.79$, 0.028, 0.79, 0.011). For detection of density change, the average confidence was 4.5 for both methods (no significant differences).

Agreement With the Standard of Reference

The observers' scores never differed more than 1 category from the standard of reference, with respect to changes of density and size of the GGOs; in other words, it never occurred that observer and standard of reference stated the opposite (eg, density increase and decrease, shrinkage and growth, respectively). The agreement of the individual observers with the standard of reference is reported in Table 1. Following the terminology of Landis and Koch[26] ($0.4 < \kappa_w \leq$ 0.6: moderate agreement; $0.6 < \kappa_w \leq 0.8$: substantial agreement;

TABLE 1. Overall Agreement With the Standard of Reference

	Observer 1	Observer 2	Observer 3	Observer 4	Mean
Size					
Standard	0.64	0.61	0.58	0.60	0.61
Subtraction	0.72 (ns)	0.70 (ns)	0.87*	0.74†	0.76
Density					
Standard	0.57	0.52	0.52	0.51	0.53
Subtraction	0.68 (ns)	0.60 (ns)	0.74†	0.54 (ns)	0.64

Agreement κ_w with the standard of reference for 59 and 58 cases for size and density assessment, respectively. The significance is given between brackets: *($P <$ 0.001) and †($P < 0.05$) indicate a significant difference in agreement κ_w of the subtraction method compared to the standard method.
ns indicates no significant difference ($P \geq 0.05$).

$\kappa_w > 0.8$: almost perfect agreement), the mean agreement with the standard of reference was substantial for size change detection, and moderate for density change detection with the standard method. The agreement increased with the use of a subtraction image, on average from 0.61 to 0.76 for size change detection, and from 0.53 to 0.64 for density change detection. For size change detection, the increase is significant for 2 raters ($P = 0.33$, 0.29, <0.001, 0.029); for density change detection, the increase is significant for one observer ($P = 0.20$, 0.44, 0.009, 0.78).

Table 2 lists the impact of image subtraction for the 2 subsets of nodules dependent on the readers' confidence.

For subgroup 1 lesions, on average, the agreement only marginally increased from 0.86 to 0.91 for size change and from 0.76 to 0.78 for density change detection. One of the observers achieved a significantly higher agreement with the reference standard with the subtraction image for the detection of size change ($P = 0.004$). For all other readers, we did not find significant differences ($P = 0.31$, 0.97, 1.00, respectively, for size change and $P = 0.73$, 0.38, 0.51, 0.69, respectively, for density change).

TABLE 2. Agreement With the Standard of Reference, Differentiation

	Observer 1	Observer 2	Observer 3	Observer 4	Mean
Subgroup 1					
Size					
Standard	0.92	0.87	0.79	0.88	0.86
Subtraction	0.85 (ns)	0.87 (ns)	0.93*	1.00 (ns)	0.91
Density					
Standard	0.72	0.72	0.75	0.83	0.76
Subtraction	0.68 (ns)	0.83 (ns)	0.81 (ns)	0.79 (ns)	0.78
Subgroup 2					
Size					
Standard	0.14	0.46	0.09	0.33	0.26
Subtraction	0.50*	0.59 (ns)	0.70†	0.51*	0.57
Density					
Standard	0.20	0.41	−0.11	0.28	0.19
Subtraction	0.61*	0.47 (ns)	0.50*	0.32 (ns)	0.47

Agreement κ_w with the standard of reference. Subgroup 1 are those cases with observer confidence 5. Subgroup 2 are the cases with confidence 1 to 4. The significance is given between brackets: *($P < 0.05$) and †($P < 0.001$) indicate a significant difference in agreement κ_w of the subtraction method compared to the standard method.
ns indicates no significant difference ($P \geq 0.05$).
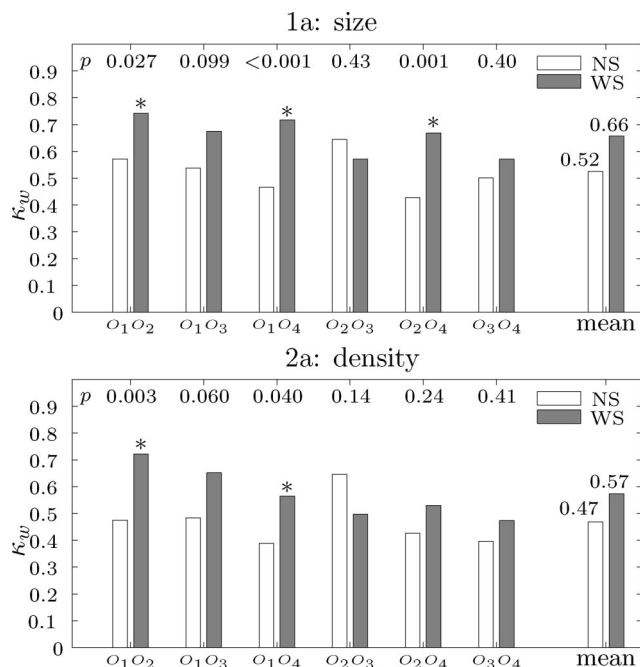
FIGURE 3. Interobserver agreement for 59 and 58 cases for size and density assessment, respectively. O_iO_j indicates the agreement between observers *i* and *j* when using a subtraction image (WS, gray bar), and separately, without using the subtraction image (NS, white bar). At the top, *P*-values are given. Statistically significant improvement of κ_w is denoted by an asterisk.

For subgroup 2 lesions, agreement with the standard of reference improved substantially with the availability of the subtraction image. On average, the agreement increased from 0.26 to 0.57 and from 0.19 to 0.47 with the use of image subtraction. The improvements were significant for 3 of the 4 observers for the detection of size change ($P = 0.010, 0.25, <0.001, 0.046$), and for 2 of the 4 observers for the detection of density change ($P = 0.002, 0.67, 0.001, 0.67$).

Interobserver Agreement

We calculated interobserver agreement separately for both reading conditions, ie, with and without availability of the subtraction image. Results are presented in Figure 3. With 4 readers, there were 6 combinations of reader pairs, for each of the 2 reading conditions. For the detection of size change, 5 of the 6 reader combinations showed an increase in agreement with the availability of subtraction image, differences were significant for 3 reader pairs ($P = 0.027, 0.099, <0.001, 0.43, 0.001, 0.40$). For the detection of density change, again 5 of the 6 reader pairs showed an increase in agreement with the availability of the subtraction image, differences were significant for 2 reader pairs ($P = 0.003, 0.060, 0.040, 0.14, 0.24, 0.41$). Both for size and density assessment, the mean agreement between observers is higher with the subtraction method, compared with the standard visual comparison method.

DISCUSSION

Although for the evaluation of solid nodules, automatic volumetry represents an established and accepted tool for lesion characterization, such methodology is still under development for part-solid and ground-glass lesions. Current clinical practice is side-by-side comparison of follow-up scans, assisted by diameter measurements,

each with its own constraints as is known from multiple studies with solid nodules.[5]

This article presents the results of a study on the impact of a subtraction image on observer performance for the detection of change of GGOs on chest CT scans. We compared current clinical practice of visual comparison of follow-up scans with readings made with additional availability of a subtraction image. Because in most of these lesions no absolute standard of reference in terms of histology or verified volumetry was available, we used the judgment of an experienced radiologist who had all clinical, reading, and imaging data available, as a standard of reference. This standard has of course its limitations but represented–in our opinion–a sufficient superiority compared with the individual readings. Lesion pairs that could not be classified with high confidence even with all information available were excluded from further data analysis. This was done to minimize the uncertainty on the data analysis, which is inevitably introduced by a nonperfect standard. Exclusion of nodules may have introduced a bias because difficult nodules are omitted from evaluation. Furthermore, exclusion of these lesions decreases the clinical applicability of the technique. Separate analysis of these cases, however, showed that it was virtually impossible for observers to determine size change, as indicated by very low weighted kappa values (≤0.19), both for interobserver variability and for agreement with the observer that established the standard of reference. This indicates that for these difficult nodules, the establishment of a meaningful standard of reference was not possible on the basis of chest CT scans. We therefore excluded them from data analysis. The lack of an absolute standard such as histology or reliable volumetry limits the clinical applicability of our study at that point. On the other hand, we conclude from the results that at least for the majority of the ground-glass nodules found in a screening setting, the proposed method was helpful for the assessment of density and volume changes and turned out to be superior to visual assessment alone. A standard of reference set by a phantom would have been an alternative, but a representative phantom with known nodule volumes is currently only available for solid nodules.[27] Comparable phantoms are not yet available for ground-glass nodules[28] because these ground-glass nodules tend to have very irregular and less well defined borders.

We found that the use of a subtraction image improves the detection of changes. When observers had a high level of confidence even without image subtraction, the agreement with the standard of reference was high, both with and without subtraction. For a lower level of confidence, image subtraction improved observer performance substantially. Taking interobserver agreement and agreement with the standard of reference as performance measures, we found that image subtraction provides a more reproducible and accurate method for detecting change in size and density of a GGO. Although we could not prove that the algorithm is applicable to all GGOs (because of the excluded cases), the results suggest that the proposed scheme leads to an improved assessment of subtle changes in GGOs, compared with visual comparison alone.

Diameter measurements of lung tumor size on CT scans are often inconsistent.[29–31] In our study, the perceived need to perform diameter measurements varied strongly between observers: some used diameter measurements in 57% (48/84), others in only 5% (4/84) of images pairs. When the subtraction image was supplied, the number of diameter measurements dropped substantially, indicating that the subtraction image replaced the perceived need for diameter measurements. Reading of the additional subtraction image had no impact on the average reading time. Although some readers needed more time for assessment of an additional image, others reached a decision faster when the subtraction image was available.

Automatic or semi-automatic volumetric or densitometric tools are also under development for GGOs, but require further study. Manual segmentation is an option but may suffer from reproducibility problems and is labor-intensive. Saito et al[32] used density slicing (the selection of voxels with values in a range of intensities) to segment a GGO. Sumikawa et al[33] employed a combination of adaptive thresholding and morphologic pruning. A more sophisticated method based on voxel classification together with a vessel suppression technique was proposed by Zhou et al.[19] Our technique provides an alternative that is based on visual image evaluation, and uses subtraction to enhance subtle differences.

Application of the subtraction method in clinical practice currently requires manual placement of an ellipsoid around the nodule. However, the only conditions are that the ellipsoid contains the nodule and excludes most of the surrounding parenchyma; thus, a rough delineation will suffice. After placement of the ellipsoid, the scans need to be registered and subtracted, which requires about 10 minutes with the current implementation. For convenient use in a clinical environment, our current implementation requires improvement of the user interface and integration of the different parts of the method into a single user interface (the placement of the ellipsoid and the registration are currently performed in separate programs). No extra expertise of the observers was required, other than a 15-minute training period to familiarize the observers with the user interface.

We conclude that image subtraction after registration for the follow-up of ground-glass opacities improves observer performance for the detection of nodule size and density change on chest CT, especially when the level of confidence of observers is low without the subtraction image, and it decreases interobserver variability.

ACKNOWLEDGMENTS

REFERENCES

1. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet*. 1999; 354:99–105.
2. Henschke CI, Yankelevitz DF, Libby DM, et al. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med*. 2006;355:1763–1771.
3. National Cancer Institute. National Lung Screening Trial. Available at: www.cancer.gov/nlst. Accessed December 27, 2007.
4. Van Iersel CA, De Koning HJ, Draisma G, et al. Risk-based selection from the general population in a screening trial: Selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). *Int J Cancer*. 2007;120:868–874.
5. Marten K, Engelke C. Computer-aided detection and automated CT volumetry of pulmonary nodules. *Eur Radiol*. 2007;17:888–901.
6. Nakata M, Sawada S, Saeki H, et al. Prospective study of thoracoscopic limited resection for ground-glass opacity selected by computed tomography. *Ann Thorac Surg*. 2003;75:1601–1605.
7. Nakajima R, Yokose T, Kakinuma R, et al. Localized pure ground-glass opacity on high-resolution CT: histologic characteristics. *J Comput Assist Tomogr*. 2002;26:323–329.
8. Suzuki K, Asamura H, Kusumoto M, et al. "Early" peripheral lung cancer: prognostic significance of ground-glass opacity on thin-section computed tomographic scan. *Ann Thorac Surg*. 2002;74:1635–1639.
9. Henschke CI, Yankelevitz DF, Mirtcheva R, et al. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *AJR Am J Roentgenol*. 2002;178:1053–1057.
10. Arenberg D. Bronchioloalveolar lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132:306–313.
11. Yang ZG, Sone S, Takashima S, et al. High-resolution CT analysis of small peripheral lung adenocarcinomas revealed on screening helical CT. *AJR Am J Roentgenol*. 2001;176:1399–1407.
12. Henschke CI, Yankelevitz DF, Smith JP, et al. Screening for lung cancer: the early lung cancer action approach. *Lung Cancer*. 2002;35:143–148.
13. Kakinuma R, Ohmatsu H, Kaneko M, et al. Progression of focal pure ground-glass opacity detected by low-dose helical computed tomography screening for lung cancer. *J Comput Assist Tomogr*. 2004;28:17–23.
14. Nakamura H, Saji H, Ogata A, et al. Lung cancer patients showing pure ground-glass opacity on computed tomography are good candidates for wedge resection. *Lung Cancer*. 2004;44:61–68.
15. Gietema HA, Schaefer-Prokop CM, Mali WPTM, et al. Pulmonary nodules: Interscan variability of semiautomated volume measurements with multisection CT–influence of inspiration level, nodule size, and segmentation performance. *Radiology*. 2007;245:888–894.
16. Wormanns D, Kohl G, Klotz E, et al. Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. *Eur Radiol*. 2004;14:86–92.
17. Kostis WJ, Yankelevitz DF, Reeves AP, et al. Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up CT. *Radiology*. 2004;231:446–452.
18. Goodman LR, Gulsun M, Washington L, et al. Inherent variability of CT lung nodule measurements in vivo using semiautomated volumetric measurements. *AJR Am J Roentgenol*. 2006;186:989–994.
19. Zhou J, Chang S, Metaxas DN, et al. Automatic detection and segmentation of ground glass opacity nodules. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*. 4190, 2006;784–791.
20. Xu D, Gietema H, de Koning H, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. *Lung Cancer*. 2006;54: 177–184.
21. Thévenaz P, Unser M. Optimization of mutual information for multiresolution image registration. *IEEE Trans Image Process*. 2000;9:2083–2099.
22. Staring M, Klein S, Pluim JPW. A rigidity penalty term for nonrigid registration. *Med Phys*. 2007;34:4098–4108.
23. Sluimer IC, Prokop M, Van Ginneken B. Toward automated segmentation of the pathological lung in CT. *IEEE Trans Med Imaging*. 2005;24:1025–1038.
24. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70:213–220.
25. Altman DG. Practical statistics for medical research. London: Chapman & Hall/CRC; 1991.
26. Landis RJ, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
27. Das M, Muhlenbruch G, Katoh M, et al. Automated volumetry of solid pulmonary nodules in a phantom: accuracy across different CT scanner technologies. *Invest Radiol*. 2007;42:297–302.
28. Multipurpose Chest Phantom. Available at: www.kyotokagaku.com/products/detail03/ph-1.html. Accessed September 9, 2008.
29. Erasmus JJ, Gladish GW, Broemeling L, et al. Interobserver and intraobserver variability in measurement of nonsmall-cell carcinoma lung lesions: implications for assessment of tumor response. *J Clin Oncol*. 2003;21:2574–2582.
30. Revel MP, Bissery A, Bienvenu M, et al. Are two-dimensional CT measurements of small noncalcified pulmonary nodules reliable? *Radiology*. 2004; 231:453–458.
31. Jennings SG, Winer-Muram HT, Tarver RD, et al. Lung tumor growth: assessment with CT–comparison of diameter and cross-sectional area with volume measurements. *Radiology*. 2004;231:866–871.
32. Saito H, Minamiya Y, Kawai H, et al. Usefulness of circumference difference for estimating the likelihood of malignancy in small solitary pulmonary nodules on CT. *Lung Cancer*. 2007;58:348–354.
33. Sumikawa H, Johkoh T, Nagareda T, et al. Pulmonary adenocarcinomas with ground-glass attenuation on thin-section CT: quantification by three-dimensional image analyzing method. *Eur J Radiol*. 2008;65:104–111.