

Coronary Artery Calcification Scoring in Low-Dose Ungated CT Screening for Lung Cancer: Interscan Agreement

Peter C. A. Jacobs^{1,2}
Ivana Isgum³
Martijn J. A. Gondrie^{1,2}
Willem P. Th. M. Mali¹
Bram van Ginneken³
Mathias Prokop¹
Yolanda van der Graaf²

Keywords: coronary artery calcification score, interscan agreement, low dose, lung cancer screening, ungated CT

DOI:10.2214/AJR.09.3047

Received May 12, 2009; accepted after revision November 5, 2009.

Supported by a program grant from The Netherlands Organization for Scientific Research—Medical Sciences (NOW-MW project no. 40-00812-98-07-005).

¹Department of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands.

²Julius Center for Health Sciences and Primary Care, Department of Radiology, University Medical Center, Universiteitsweg 100, PO Box 85060, 3508 GA, Utrecht, The Netherlands. Address correspondence to P. C. A. Jacobs (p.c.a.jacobs@umcutrecht.nl).

³Image Science Institute, University Medical Center Utrecht, Utrecht, The Netherlands.

AJR 2010; 194:1244–1249

0361–803X/10/1945–1244

© American Roentgen Ray Society

OBJECTIVE. In previous studies detection of coronary artery calcification (CAC) with low-dose ungated MDCT performed for lung cancer screening has been compared with detection with cardiac CT. We evaluated the interscan agreement of CAC scores from two consecutive low-dose ungated MDCT examinations.

SUBJECTS AND METHODS. The subjects were 584 participants in the screening segment of a lung cancer screening trial who underwent two low-dose ungated MDCT examinations within 4 months (mean, 3.1 ± 0.6 months) of a baseline CT examination. Agatston score, volume score, and calcium mass score were measured by two observers. Interscan agreement of stratification of participants into four Agatston score risk categories (0, 1–100, 101–400, > 400) was assessed with kappa values. Interscan variability and 95% repeatability limits were calculated for all three calcium measures and compared by repeated measures analysis of variance.

RESULTS. An Agatston score > 0 was detected in 443 baseline CT examinations (75.8%). Interscan agreement of the four risk categories was good ($\kappa = 0.67$). The Agatston scores were in the same risk category in both examinations in 440 cases (75.3%); 578 participants (99.0%) had scores differing a maximum of one category. Furthermore, mean interscan variability ranged from 61% for calcium volume score to 71% for Agatston score ($p < 0.01$). A limitation of this study was that no comparison of CAC scores between low-dose ungated CT and the reference standard ECG-gated CT was performed.

CONCLUSION. Cardiovascular disease risk stratification with low-dose ungated MDCT is feasible and has good interscan agreement of stratification of participants into Agatston score risk categories. High mean interscan variability precludes the use of this technique for monitoring CAC scores for individual patients.

Because prolonged and heavy smoking is a strong etiologic factor in the development of lung cancer and cardiovascular disease (CVD), simultaneous screening for both diseases in a high-risk population of heavy smokers and former smokers is an attractive strategy for maximizing the beneficial effects on survival in lung cancer screening programs and minimizing the radiation dose for individuals in the screening population. The presence and extent of coronary artery calcification (CAC) detected with electron-beam CT and ECG-gated MDCT have been found to be independent predictors of CVD events [1]. By contrast, screening for lung cancer is performed with low-dose ungated MDCT scanning protocols. Although the clinical importance of CAC detected with both ungated MDCT [2] and low-dose CT [3,

4] has been established, factors negatively influencing variability in calcium-scoring CT, such as image noise [5], intraobserver and interobserver variability [6], motion artifacts [7], and partial volume effects [8], have always been deemed to have too great an influence on ungated low-dose MDCT to consider it a reasonable alternative. However, two studies [9, 10] comparing calcium measurements in low-dose ungated MDCT with results from ECG-gated MDCT concluded that comparable results can be obtained.

To be truly useful as a screening tool for cardiovascular disease, low-dose ungated MDCT should also show sufficient interscan agreement. Even in MDCT performed specifically for calcium scoring, substantial variability occurs in repeated scans of the same person [11]. However, if calcium scoring in low-dose ungated MDCT is used as a marker

Calcium Scoring and CT for Lung Cancer

for predicting CVD events, the main concern should be not the absolute or relative interscan variability of calcium scores but correct stratification of patients into CVD risk categories, such as the system based on the Agatston score described by Rumberger et al. [12]. Those authors suggested clinically applicable cutoffs in the continuous Agatston score (0, low risk; 1–100, mild risk; 100–400, moderate risk; and > 400, high risk) that correspond to increasing levels of risk of coronary events. In analogy to application of the Framingham risk score, patients in the highest risk category (> 400) need aggressive preventive treatment of CVD risk factors, and this recommendation should possibly be extended to patients in the intermediate risk category (100–400). The main purpose of our study was to determine the interscan agreement of Agatston score risk categories in repeated low-dose ungated MDCT in an evaluation of the usefulness of this technique for prediction of CVD events in participants in a lung cancer screening trial.

Subjects and Methods

Approval was obtained from the institutional review committees of all participating study sites. Informed consent was obtained from all participants.

Participants

A search of population registries yielded the cases of 15,822 subjects between 50 and 75 years of age who were recruited to participate in the Nederlands-Leuvens Longkanker Screenings Onderzoek (NELSON) study, a lung cancer screening trial performed in four regions in The Netherlands and Belgium. The inclusion criteria were, first, that participant was a current or former smoker with a smoking history of more than 15 cigarettes per day for more than 25 years or more than 10 cigarettes per day for more than 30 years who stopped smoking more than 10 years before study enrollment and, second, that the participant be able to climb two or more flights of stairs. A detailed description of patient selection for the study has been published [13].

Between April 2004 and July 2006, 7,557 participants randomized to the screening group at four participating medical centers underwent a baseline CT examination. For the current study, a subgroup of 584 participants (497 men, 87 women; mean age, 59.9 ± 5.8 [SD] years) from one of the four participating centers was identified who underwent repeated scanning within 4 months of the baseline CT scan. The indication for repetition of CT was detection of a pulmonary nodule measuring 50–500 mm³ at baseline CT [14].

The mean period between the baseline and repeated scans was 3.1 ± 0.6 months (range, 2.0–4.0 months). That only limited progression of CAC can be expected in this short interval [15] made this opportunity a good one for investigating the interscan variability of CAC measurements.

Low-Dose Chest CT Protocol

All baseline and repeated CT scans performed at the participating study site were conducted with a 16-MDCT scanner (Mx8000 IDT, Philips Healthcare). Scanning was performed in helical mode with 16×0.75 mm collimation and pitch of 1.3–1.5. The scanning parameters have been described extensively [14]. Transverse images with 1.0-mm section thickness and a 0.7-mm increment were acquired from the level of the lung bases to the lung apices. The smallest field of view was chosen to include the outer rib margins at the widest dimension of the chest.

Acquisition was performed at suspended maximal inspiration after the participant received instructions about breath-holding. Total acquisition duration was a single breath-hold (≈ 10 seconds). No ECG triggering was performed; no contrast agent was administered. For the analysis of CAC, the raw data were reconstructed into 3.1-mm overlapping sections with a 1.4-mm increment. Low-dose exposure settings were applied according to body weight: 30 mAs at a tube voltage of 120 kVp for participants weighing 80 kg or less and 30 mAs at a tube voltage of 140 kVp for subjects weighing more than 80 kg. Radiation exposure (volume CT dose index) was 2.2 mGy for the lighter and 3.5 mGy for the heavier participants.

CAC Assessment

All CAC scoring was performed with software described by Isgum et al. [16]. Two readers, a research physician with 2 years of experience and a radiologist with 3 years of experience in cardiac CT, blinded to the participant's age, sex, and name, independently read the images from all 1,168 CT scans. Images from the two scans of a single participant were read by one observer at least 2 weeks apart to eliminate the effect of interobserver variability. Scan pairs were divided equally between the two readers. Before the study, both readers reviewed a training set of images from 50 randomly selected CT scans in the NELSON study database to assess interobserver variability (intraclass $R = 0.97$).

Calcium scoring software written in C++ was used to define calcified plaque as all regions of three or more adjoining voxels (0.7 mm³) with attenuation greater than 130 HU. An investigator manually identified a point in each calcified lesion. Three-dimensional component labeling with 26-connectivity was automatically performed to

mark all connected voxels as calcification. Care was taken not to include noncoronary calcifications (e.g., valve calcifications) and hyperattenuating foci due to image noise [16, 17]. Total Agatston score, calcium volume score, and calcium hydroxyapatite mass score were output by the software program as outlined by Ulzheimer and Kalender [18]. Separate scores were calculated for the left main, left anterior descending, left circumflex, and right coronary arteries. Agatston scores were further categorized into four groups (0, 1–100, 101–400, and > 400) to be used for CVD risk stratification as outlined by Rumberger et al. [12].

Statistical Analysis

Intraclass correlation coefficients, kappa values, and variability were used to express interobserver and interscan agreement of CAC scores. First, the presence of CAC (Agatston score > 0) on scan 1 versus the presence of calcification on scan 2 was assessed for all participants ($n = 584$). Disagreement between the presence of CAC on two scans of the same patient was counted as a discordant pair. Agreement on the presence (yes or no) of calcium was calculated with kappa statistics.

All participants were additionally stratified into a CVD risk category for scan 1 and for scan 2 on the basis of Agatston score: 0, low risk; 1–100, moderate risk; 101–400, intermediate risk; and > 400, high risk [12]. Agreement of risk category stratification also was assessed with kappa statistics. Relative interscan variability of two CAC scores was calculated as: absolute (score 1 – score 2) / mean (score 1 + score 2) $\times 100\%$. In participants with an Agatston score > 0 in at least one scan ($n = 461$), we used the regression method for nonuniform differences to establish 95% repeatability limits for all three measures [19]. In this method, the absolute interscan difference (D) of each calcium measure was linearly modeled against the mean calcium measures (M). All models were adjusted for the total amount of calcium, measured according to the mean (natural log transformed) calcium score. We chose to force these models through the origin because the interscan difference is zero when the mean calcium measure is zero [20]. The resulting regression line is $D = b_1M$, where b_1 is the slope of the line.

We then modeled the absolute values of the unstandardized residuals ($|R|$) from the previous regression models against the mean calcium measures (M): $R = c_0 + c_1M$, where c_0 is the intercept and c_1 is the slope of the regression line. The 95% repeatability interval can be calculated by combining the two regression equations as follows: $b_1M \pm 2.46(c_0 + c_1M)$. On the assumption that the systematic difference between the two scans (b_1M) equals zero, this equation can be written as

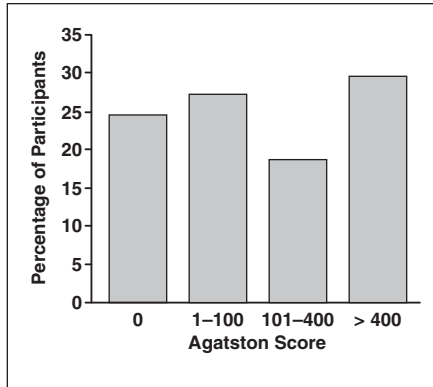


Fig. 1—Frequency histogram shows distribution of Agatston scores from baseline CT scans of 584 subjects.

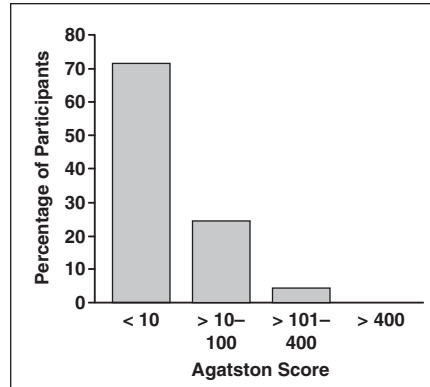


Fig. 2—Bar chart shows distribution of Agatston scores among 49 participants with coronary artery calcium in only one of two scans (discordant pairs).

TABLE 1: Risk Stratification of Subjects by Agatston Score Risk Category From Baseline and Repeated Low-Dose Ungated MDCT

Baseline Score	Score at Repeated MDCT				Total
	0	1-100	101-400	> 400	
0	125	18	0	0	143
1-100	29	102	26	2	159
101-400	2	22	59	27	110
> 400	0	4	14	154	172
Total	156	146	99	183	584

$2.46(c_0 + c_1M)$. In this method, 2.46 is substituted for 1.96 because it is assumed that the absolute values of the residuals follow a half-normal distribution. It therefore is necessary to multiply 1.96 by $\sqrt{\pi} / 2$ [19]. Absolute interscan difference was plotted against mean calcium score with the Bland-Altman approach. The equation $2.46(c_0 + c_1M)$ can be used to calculate the upper and lower 95% repeatability limits for the absolute interscan difference of a mean calcium score value. Repeated measures analysis of variance was used to assess differences in relative interscan differences in the three calcium measures. Wilcoxon's tests were used for post hoc comparisons and p for significance modified by Bonferroni adjustment ($p < 0.018$).

Results

A total of 584 participants (497 men, 87 women; mean age, 59.9 ± 5.8 years) were included in this study. Figure 1 shows the frequency distribution of CAC on the baseline scans of all 584 participants subdivided into Agatston score-based risk categories. A total of 443 participants (75.8%) had any CAC (Agatston score > 0) detected on the baseline CT scan. The median Agatston score at baseline was 80.3 (range, 0-9,596) and at repeated CT was 78.1 (range, 0-7,659). Agreement

between absolute Agatston scores on two low-dose ungated CT scans was an intraclass correlation coefficient of 0.94.

Overall agreement between scans for the presence (yes or no) of any CAC was 91.6% (535 of 584 participants). Only 49 pairs of scans (8.4%) were discordant, resulting in a good kappa value of 0.78. Figure 2 is a bar chart of the distribution of absolute Agatston scores of the 49 participants with CAC in only one scan (discordant pairs). Thirty-five of these participants (71.4%) had an Agatston score < 10 , which was negligible and was likely caused by minor motion artifacts. Only two participants (4.1%) had Agatston scores that were substantially out of range (maximum, 278.7) owing to major motion artifacts.

Table 1 shows the interscan agreement of Agatston scores on two low-dose CT scans of the same patient for the four categories used in CVD risk stratification. In approximately three of four participants (440 of 584, 75.3%), no shift in Agatston score risk category occurred between scan 1 and scan 2. The unweighted kappa statistic showed good agreement ($\kappa = 0.67$) [21]. A shift of more than one category was found in only eight of the participants (1.4%).

Figure 3 shows standard Bland-Altman plots in which the means of calcium scores from scan 1 and scan 2 are plotted against the adjusted absolute interscan difference. As expected, all three plots show a nonuniform relation between the extent of CAC and the measurement error; that is, interscan variability increased as the total amount of calcium increased. Standard 95% CIs (calculated as the mean difference $\pm 1.96 \times SD$) do not correspond well with this type of relation. Therefore, 95% confidence limits were calculated with nonuniform regression analysis and are represented by lines on the plots. The slopes of these lines (r) can be interpreted as a measure of reproducibility. The steeper the slope of this line (higher value of r), the less reproducible is the calcium measure. The vertical distance between the lines indicates the measurement error (95% CI) for a given mean calcium score value; that is, 95% of the time the absolute interscan difference for a given calcium score will fall within these limits. In this study, calcium volume score ($r = 0.5175$) had better reproducibility than Agatston score ($r = 0.5985$) and calcium mass score ($r = 0.5696$). Table 2 shows the relative interscan variability for all three calcium measures in participants who had any CAC on at least one of two CT scans ($n = 461$). A statistically significant difference ($p < 0.01$) was found in pairwise comparison of the means of variability of the three calcium measures.

Discussion

We found that use of a low-dose ungated MDCT technique for detection and quantification of CAC in a lung cancer screening trial gives good interscan agreement in the assignment of participants to Agatston score categories for CVD risk stratification. These results support the idea that CAC scoring as part of low-dose ungated MDCT can be a useful tool for assessing the risk of CVD among persons undergoing lung cancer screening.

Our results provide incremental evidence to the conclusions of two previous studies [9, 10] of this issue. Both studies compared a low-dose ungated MDCT protocol and an ECG-gated CT protocol (the reference standard for CAC scoring) with respect to accuracy in detection and categorization of CAC. Using a 40-MDCT unit, Kim et al. [9] found concordance of 83% ($n = 106$) in stratifying participants into the same risk category with low-dose CT compared with ECG-gated CT and a maximum difference of only one category in the other 12 participants. Wu et al.

Calcium Scoring and CT for Lung Cancer

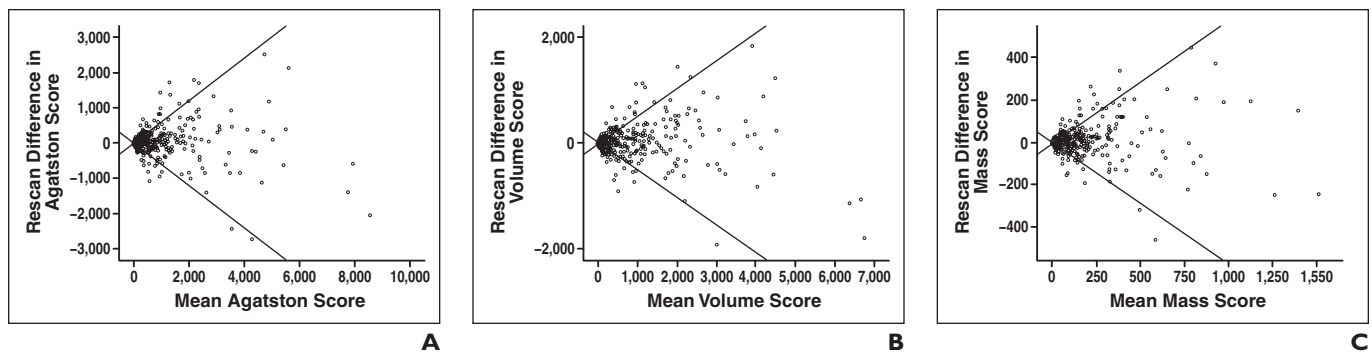


Fig. 3—Bland-Altman plots show absolute interscan difference plotted against means of three calcium measures with 95% repeatability limits adjusted for coronary artery calcification.

A, Agatston score, $r = 0.5985$.

B, Calcium volume score, $r = 0.5175$.

C, Calcium mass score, $r = 0.5696$.

TABLE 2: Interscan Variability of Coronary Artery Calcification in Participants With Agatston Score > 0 on At Least One Scan ($n = 461$)

Value	Agatston Score	Volume Score	Mass Score
Intraclass correlation coefficient	0.94	0.95	0.94
Mean variability (%)	71	61	65
Median variability (%)	51	36	43

Note—Post hoc comparisons after a repeated measures analysis of variance showed statistically significant ($p < 0.01$) differences in mean variability of all three calcium measures.

[10], using a 16-MDCT unit, found concordance of 93% ($n = 450$) compared with the reference standard.

Apart from good agreement with ECG-gated MDCT, it is important to find reliably good agreement between CAC scores on repeated CT scans of the same patient. Although mean variability between scores on both CT examinations was as high as 60–70%, we found that 75% of participants were placed in the same risk category on two low-dose ungated MDCT scans ($\kappa = 0.67$). Taken together, these results support the idea that CAC scoring of absolute calcium scores in low-dose ungated MDCT may not be highly accurate but can be used reliably for CVD risk stratification.

An important difference between the two previous studies and ours is the racial makeup of the study populations. The two previous studies [9, 10] were conducted with self-referred Asian populations, resulting in a low prevalence of CAC compared with our study, which was conducted with a predominantly white population. Because variability in scores (whether between or within techniques) is strongly related to the total amount of calcification that can be scored in an individual, it is an important strength of this study that comparable results were found in a population with a far higher prevalence of CAC.

In an earlier study, Shemesh et al. [22] used visual (semiquantitative, grade 0–12) grading of CAC to determine CAC scores in 4,250 participants in a lung cancer screening study. They concluded that with this tool, CAC scoring can be used for CVD risk stratification. Their choice to perform visual grading might have stemmed from the fact that their study was performed with a 4-MDCT unit, resulting in less spatial and temporal resolution than can be achieved with a 16-MDCT unit. Semiautomated scoring has the advantage over visual grading with respect to interobserver and intraobserver reproducibility. The subjective assessment (size, attenuation) of a vascular calcification is replaced by a software tool, with which calcifications are identified on the basis of a clear attenuation threshold (130 HU) and calcium scores are automatically calculated.

Apart from interscan agreement of Agatston risk categories, we investigated the utility of low-dose ungated MDCT in detection of the presence (yes or no) of CAC between two examinations. A previous report [23] showed that the absence of any CAC, though not completely excluding the risk of coronary artery disease, is an important indicator of absence of significant (> 50%) stenosis of the coronary arteries. Therefore, a high negative predictive value for interscan com-

parisons is desirable. The rate of occurrence of discordant pairs (calcium disappearing or suddenly appearing between two examinations) in our study was 8%. The rates of reporting of discordant pairs in previous studies of interscan agreement performed with ECG gating or triggering and varied tube currents have ranged from 0% to 6% [3, 17, 24, 25]. The result in our study is only moderately worse than these rates of discordant pairs and was obtained at a considerably reduced radiation dose to each participant. In most cases, the higher noise levels in low-dose ungated MDCT than in ECG-gated MDCT are the cause of discordance [5]. Use of a body weight–adapted scanning protocol for low-dose ungated MDCT may improve noise levels [26]. Even at present, however, the absence or presence of CAC can be established reliably in 92% of subjects.

With respect to interscan variability, recent studies [3, 25, 27] conducted with ECG-gated MDCT have shown interscan variability ranging from 12% to 32%, and studies [5, 8, 28–32] of electron-beam CT conducted in the late 1990s showed variability ranging from 13% to 51% for both Agatston and calcium volume score algorithms. We have found interscan variability of 60–70%. Therefore, we have to conclude that ECG-gated techniques remain the reference standard for obtaining accurate CAC scores in individual patients and monitoring of CAC over time, but low-dose ungated CT can be used for adequate CVD risk stratification in screening populations.

Although percentage differences are the most common way to present variability, we emphasize that this measure is inappropriate for inferring true variability. Only absolute differences in scores reflect the true variability (e.g., an absolute interscan difference of only 8

points of Agatston score can be presented as a relative difference as high as 200% (Agatston score scan 1, 0; Agatston score scan 2, 8) or as little as 2% (Agatston score scan 1, 400; Agatston score scan 2, 408). In this study, calcium volume and calcium mass scores had significantly better interscan agreement than did Agatston score. This finding is in accordance with results of previous studies [33]. Rather surprisingly, volume score performs significantly better than mass score. Results of a study by Hoffmann et al. [34] suggested the superiority of the mass score. The theoretic advantage of mass score over volume score is that density information is used to correct for partial volume effects. In our ungated scans, however, the influence of motion artifacts on the density of calcified plaques was far greater than the influence of partial volume effects. We postulate, therefore, that the positive effect of the calcium mass score is outdone by the relatively poor quality of the scans.

One limitation of our study was that we calculated interscan variability between two scans at a mean interval of 3 months. To a certain extent, this practice interfered with a one-to-one comparison with other studies, in which two scans usually have been performed only minutes apart. Normal progression of CAC is estimated at 14–27% per year [15]. Consequently, part of the variability observed in our study may be attributed to the real progression of CAC over the course of 3 months. Therefore, we expect that the results of this study would have been even better if we had performed two baseline scans for all participants. Another limitation of our study was that in establishing the 95% confidence limits for the three calcium measures, we might have controlled for participant-specific covariates (body mass index, CAC score) to improve generalization of these results. However, we could not adjust calcium scores for body mass index. Previous work has shown that greater body mass index is associated with lower interscan reproducibility, possibly caused by an increase in image noise [35]. The use of a low-dose scan protocol for overweight participants with low CAC scores is likely to interfere with accurate detection of true calcification. Even when it would have been possible to control for all participant-specific covariates, however, use of different scanner types from different vendors seems to limit even more profoundly the generalization of 95% confidence limits derived from a single study [24].

We conclude that CAC scoring with low-dose ungated MDCT as part of lung cancer

screening is reliable and has good interscan agreement for stratification of participants into CVD risk categories. This capability makes low-dose ungated MDCT a potentially valuable tool in the assessment of cardiovascular risk in large screening populations at a substantially reduced radiation dose.

References

- Greenland P, Bonow RO, Brundage BH, et al. ACCF/AHA 2007 clinical expert consensus document on coronary artery calcium scoring by computed tomography in global cardiovascular risk assessment and in evaluation of patients with chest pain: a report of the American College of Cardiology Foundation Clinical Expert Consensus Task Force (ACCF/AHA Writing Committee to Update the 2000 Expert Consensus Document on Electron Beam Computed Tomography) developed in collaboration with the Society of Atherosclerosis Imaging and Prevention and the Society of Cardiovascular Computed Tomography. *J Am Coll Cardiol* 2007; 49:378–402
- Schmermund A, Erbel R, Silber S. Age and gender distribution of coronary artery calcium measured by four-slice computed tomography in 2,030 persons with no symptoms of coronary artery disease. *Am J Cardiol* 2002; 90:168–173
- Horiguchi J, Yamamoto H, Hirai N, et al. Variability of repeated coronary artery calcium measurements on low-dose ECG-gated 16-MDCT. *AJR* 2006; 187:118; [web]W1–W6
- Shemesh J, Evron R, Koren-Morag N, et al. Coronary artery calcium measurement with multi-detector row CT and low radiation dose: comparison between 55 and 165 mAs. *Radiology* 2005; 236:810–814
- Bielak LF, Kaufmann RB, Moll PP, McCollough CH, Schwartz RS, Sheedy PF. Small lesions in the heart identified at electron beam CT: calcification or noise? *Radiology* 1994; 192:631–636
- Kaufmann RB, Sheedy PF, Breen JF, et al. Detection of heart calcification with electron beam CT: interobserver and intraobserver reliability for scoring quantification. *Radiology* 1994; 190:347–352
- Horiguchi J, Fukuda H, Yamamoto H, et al. The impact of motion artifacts on the reproducibility of repeated coronary artery calcium measurements. *Eur Radiol* 2007; 17:81–86
- Kajinami K, Seki H, Takekoshi N, Mabuchi H. Quantification of coronary artery calcification using ultrafast computed tomography: reproducibility of measurements. *Coron Artery Dis* 1993; 4:1103–1108
- Kim SM, Chung MJ, Lee KS, Choe YH, Yi CA, Choe BK. Coronary calcium screening using low-dose lung cancer screening: effectiveness of MDCT with retrospective reconstruction. *AJR* 2008; 190:917–922
- Wu MT, Yang P, Huang YL, et al. Coronary arterial calcification on low-dose ungated MDCT for lung cancer screening: concordance study with dedicated cardiac CT. *AJR* 2008; 190:923–928
- Chen J, Krumholz HM. How useful is computed tomography for screening for coronary artery disease? Screening for coronary artery disease with electron-beam computed tomography is not useful. *Circulation* 2006; 113:125–146
- Rumberger JA, Brundage BH, Rader DJ, Kondos G. Electron beam computed tomographic coronary calcium scanning: a review and guidelines for use in asymptomatic persons. *Mayo Clin Proc* 1999; 74:243–252
- 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
- Xu DM, Gietema H, de Koning HJ, et al. Nodule management protocol of the NELSON randomised lung cancer screening trial. *Lung Cancer* 2006; 54:177–184
- Maher JE, Bielak LF, Raz JA, Sheedy PF, Schwartz RS, Peyser PA. Progression of coronary artery calcification: a pilot study. *Mayo Clin Proc* 1999; 74:347–355
- Isgum I, Rutten A, Prokop M, van Ginneken B. Detection of coronary calcifications from computed tomography scans for automated risk assessment of coronary artery disease. *Med Phys* 2007; 34:1450–1461
- Rutten A, Isgum I, Prokop M. Coronary calcification: effect of small variation of scan starting position on Agatston, volume, and mass scores. *Radiology* 2008; 246:90–98
- Ulzheimer S, Kalender WA. Assessment of calcium scoring performance in cardiac computed tomography. *Eur Radiol* 2003; 13:484–497
- Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8:135–160
- Detrano RC, Anderson M, Nelson J, et al. Coronary calcium measurements: effect of CT scanner type and calcium measure on rescan reproducibility—MESA study. *Radiology* 2005; 236:477–484
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–174
- Shemesh J, Henschke CI, Farooqi A, et al. Frequency of coronary artery calcification on low-dose computed tomography screening for lung cancer. *Clin Imaging* 2006; 30:181–185
- Rubinshtein R, Gaspar T, Halon DA, Goldstein J, Peled N, Lewis BS. Prevalence and extent of obstructive coronary artery disease in patients with

Calcium Scoring and CT for Lung Cancer

- zero or low calcium score undergoing 64-slice cardiac multidetector computed tomography for evaluation of a chest pain syndrome. *Am J Cardiol* 2007; 99:472–475
24. Chung H, McClelland RL, Katz R, Carr JJ, Budoff MJ. Repeatability limits for measurement of coronary artery calcified plaque with cardiac CT in the Multi-Ethnic Study of Atherosclerosis. *AJR* 2008; 190:321; [web]W87–W92
25. Takahashi N, Bae KT. Quantification of coronary artery calcium with multi-detector row CT: assessing interscan variability with different tube currents pilot study. *Radiology* 2003; 228:101–106
26. Mahnken AH, Wildberger JE, Simon J, et al. Detection of coronary calcifications: feasibility of dose reduction with a body weight-adapted examination protocol. *AJR* 2003; 181:533–538
27. Detrano R, Kang X, Mahaisavariya P, et al. Accuracy of quantifying coronary hydroxyapatite with electron beam tomography. *Invest Radiol* 1994; 29:733–738
28. Callister TQ, Coil B, Raya SP, Lippolis NJ, Russo DJ, Raggi P. Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam CT volumetric method. *Radiology* 1998; 208:807–814
29. Devries S, Wolfkiel C, Shah V, Chomka E, Rich S. Reproducibility of the measurement of coronary calcium with ultrafast computed tomography. *Am J Cardiol* 1995; 75:973–975
30. Mao S, Bakhsheshi H, Lu B, Liu SC, Oudiz RJ, Budoff MJ. Effect of electrocardiogram triggering on reproducibility of coronary artery calcium scoring. *Radiology* 2001; 220:707–711
31. Mohlenkamp S, Behrenbeck TR, Pump H, et al. Reproducibility of two coronary calcium quantification algorithms in patients with different degrees of calcification. *Int J Cardiovasc Imaging* 2001; 17:133–142
32. Wang S, Detrano RC, Secci A, et al. Detection of coronary calcification with electron-beam computed tomography: evaluation of interexamination reproducibility and comparison of three image-acquisition protocols. *Am Heart J* 1996; 132:550–558
33. Hong C, Bae KT, Pilgram TK. Coronary artery calcium: accuracy and reproducibility of measurements with multi-detector row CT—assessment of effects of different thresholds and quantification methods. *Radiology* 2003; 227:795–801
34. Hoffmann U, Siebert U, Bull-Stewart A, et al. Evidence for lower variability of coronary artery calcium mineral mass measurements by multi-detector computed tomography in a community-based cohort—consequences for progression studies. *Eur J Radiol* 2006; 57:396–402
35. Sevrukov A, Prapat A, Doss C, Jelnin V, Hoff JA, Kondos GT. Electron beam tomography imaging of coronary calcium: the effect of body mass index on radiologic noise. *J Comput Assist Tomogr* 2002; 26:592–597

FOR YOUR INFORMATION

Mark your calendar for the following ARRS annual meetings:

May 1–6, 2011—Hyatt Regency Chicago, Chicago, IL

April 29–May 4, 2012—Vancouver Convention Center, Vancouver, BC, Canada

April 14–April 19, 2013—Marriott Wardman Park Hotel, Washington, DC