Agreement of Visual Estimation of Coronary Artery Calcium From Low-Dose CT Attenuation Correction Scans in Hybrid PET/CT and SPECT/CT With Standard Agatston Score

Andrew J. Einstein, MD, PhD,* Lynne L. Johnson, MD,* Sabahat Bokhari, MD,* Jessica Son, MD,* Randall C. Thompson, MD,‡ Timothy M. Bateman, MD,‡ Sean W. Hayes, MD,§ Daniel S. Berman, MD§

New York, New York; Kansas City, Missouri; and Los Angeles, California

Objectives
We sought to evaluate the accuracy and reproducibility of visual estimation of coronary artery calcium (CAC) from computed tomography attenuation correction (CTAC) scans performed for hybrid positron emission tomography (PET)/computed tomography (CT) and single-photon emission computed tomography (SPECT)/CT myocardial perfusion imaging (MPI).

Background
At the time of MPI, hybrid systems obtain a low-dose, non-electrocardiogram (ECG)-gated CT scan that is used to perform attenuation correction. Utility of this CTAC scan in estimating actual CAC as measured by Agatston score (AS) on standard ECG-gated scans has not been previously studied.

Methods
A total of 492 patients, from 3 centers, receiving both MPI with CTAC and a standard CAC scan were studied. At each site, experienced readers blinded to AS reviewed CTAC images, visually estimating CAC on a 6-level scale: classifying patients as estimated AS of 0, 1 to 9, 10 to 99, 100 to 300, 400 to 999, or ≥1,000. Agreement between visually estimated coronary artery calcium (VECAC) on CTAC and AS, measured standardly and converted to the same scale, was evaluated, as was inter-reader agreement.

Results
Although CTAC images are low dose and nongated, a high degree of association was observed between VECAC and AS, with 63% of VECACs in the same category as the AS category and 93% within 1 category. Weighted kappa was 0.89 (95% confidence interval: 0.88 to 0.91, p < 0.0001). High weighted kappa statistics were observed for each site, scanner type, and sex. Readers reported identical scores in 65% of cases and scores within 1 category in 93%.

Conclusions
CAC can be visually assessed from low-dose CTAC scans with high agreement with AS. CTAC scans should be routinely assessed for VECAC. (J Am Coll Cardiol 2010;56:1914–21) © 2010 by the American College of Cardiology Foundation

In recent years, positron emission tomography (PET) and single-photon emission computed tomography (SPECT) cameras that incorporate multidetector row computed tomography (CT) scanners have been introduced into clinical practice. At the time of myocardial perfusion imaging (MPI), such hybrid systems obtain a low-dose, non-electrocardiogram (ECG)-gated CT scan that is used to perform attenuation correction for the radionuclide images (Figs. 1 and 2). Computed tomography attenuation correction (CTAC) provides the potential to eliminate false positive perfusion defects caused by soft-tissue attenuating structures such as the diaphragm and breasts (1).

Many laboratories routinely perform in the same session as MPI an additional high-resolution, ECG-gated cardiac CT scan to quantitate coronary artery calcium (CAC).
Calcium scoring provides supplementary information to MPI by identifying and assessing the extent of calcified plaques of the coronary arteries, even in the absence of the flow-limiting disease associated with perfusion defects. However, calcium scoring is not routinely performed with MPI and, when done, adds modestly to the total effective dose of radiation.

Visual comparison of nongated CTAC scans with ECG-gated scans performed in the same patients for calcium scoring suggests that much of the information relating to the extent of CAC is available from the low radiation dose, nongated CTAC scan (Figs. 1 and 2). However, the nongated scan is subject to blurring of coronary arteries, which may change CT number in some voxels, which can affect calcium scores. Moreover, lack of ECG gating has potential to introduce gaps and/or duplications in the image dataset due to cardiac motion. The purpose of this study was to evaluate the accuracy of visual estimation of CAC from CTAC scans, in comparison to “gold-standard” Agatston score (AS) measurement on standard ECG-gated CT scans, as well as to assess the reproducibility of this visual estimation.

Methods

Study design. A total of 492 patients from 3 centers performing hybrid PET/CT or SPECT/CT were studied. Institutional review board approval or waiver was obtained at each site. Each patient had received both an MPI study with CTAC and a separate CAC scan performed on the same day. Each CTAC scan was visually assessed by 2 independent readers at each site and graded on a 6-point scale, to estimate the extent of CAC. AS was determined for each CAC scan and reclassified on the same scale. Agreement between visually estimated coronary artery calcium (VECAC) scores and AS was determined using weighted kappa statistics and percentage agreements, as was inter-reader reproducibility of VECAC scores.

PET/CT and SPECT/CT scanning. At Columbia University Medical Center (116 women, 81 men), each patient was scanned using a Philips Precedence 16P SPECT/CT scanner (Philips, Cleveland, Ohio). CTAC scans were performed free-breathing.
without ECG gating, in spiral mode with pitch 0.94, collimation $16 \times 1.5$ mm, scan length 14 mm, tube voltage 120 kVp, and effective mAs 50, adjusted by the technologist according to patient habitus. Typical dose-length-product was 75 mGy·cm, corresponding to an estimated effective dose of 1.3 mSv (2). CAC scans were performed using an end-inspiratory breath-hold with ECG gating, axial mode, collimation $8 \times 3.0$ mm, scan length 14 cm, tube voltage 120 kVp, and effective mAs 70, on rare occasions increased by the technologist to reflect patient body habitus. Typical dose-length-product was 58 mGy·cm, corresponding to an estimated effective dose of 1.0 mSv.

At Mid America Heart (111 women, 75 men), each patient was scanned using a Siemens Biograph 16 PET/CT scanner (Siemens, Munich, Germany). Post-stress CTAC scans were performed using a light end-expiratory breath-hold without ECG gating, spiral mode with pitch 2, collimation $16 \times 0.75$ mm, tube voltage 120 kVp, effective mAs 9, and typical dose-length-product 15 mGy·cm, corresponding to an estimated effective dose of 0.3 mSv. CAC scans were performed using an extended craniocaudal scan also used for rest attenuation correction. This used an end-expiratory breath-hold with ECG gating, spiral mode with pitch 0.28, collimation $16 \times 1.5$ mm, tube voltage 120 kVp, effective mAs ~220 mAs adjusted by the technologist to reflect patient habitus, and typical dose-length-product 278 mGy·cm, corresponding to an estimated effective dose of 4.7 mSv.

At Cedars-Sinai (50 women, 59 men), each patient was scanned using a Siemens Biograph 64 PET/CT scanner. CTAC scans were performed free breathing without ECG gating, spiral mode with pitch 1.5, collimation $24 \times 1.2$ mm, tube voltage 120 kVp, effective mAs 11, and typical dose-length product 16 mGy·cm, corresponding to an estimated effective dose of 0.3 mSv. CAC scans were performed using an end-inspiratory breath-hold with prospective ECG gating, collimation $30 \times 0.6$ mm, tube voltage 120 kVp, effective mAs 150, and typical dose-length-product 173 mGy·cm, corresponding to an estimated effective dose of 2.9 mSv.

**Analysis of CT scans.** At each site, AS was determined using standard methodology on a dedicated workstation. At each site, 2 experienced readers blinded to the AS independently reviewed CTAC images and visually estimated CAC...
on a 6-level scale, classifying patients as having an estimated AS of 0, 1 to 9, 10 to 99, 100 to 300, 400 to 999, or ≥1,000. In addition, a single reader from the Mid America Heart and from the Cedars-Sinai sites each blindly visually estimated CAC on 99 scans from Columbia University Medical Center. For the first 20 cases, readers were given post-hoc feedback as to the correct classification based on the AS. Each reader also visually evaluated each CTAC scan for the presence or absence of calcium in each of the major coronary arteries (left main, left anterior descending [LAD], circumflex, and right coronary arteries).

Since systematically missing a proximal LAD lesion may have important clinical implications, a subsequent analysis was performed in 1 center (Columbia) of cases with a discrepancy between visual assessment of LAD calcium on CTAC images, by either reader, and its measurement (0 or ≥0) on the CAC images, to determine the location in the vessel of the erroneous visual estimation. This was performed by a single reader who visually inspected the LAD of discrepant cases and classified calcium as being present or absent in the proximal, mid, and/or distal vessel. Analysis was performed on separate days for the CTAC and CAC images, with images being presented in random order on each day, and the reader blinded to classifications from previous reading sessions.

**Effect on interpretation of borderline MPI scans.** One use of the AS in hybrid imaging can be to provide an additional source of information to assist in interpretation of borderline MPI scans. For example, for a borderline, nonextensive perfusion defect, some readers will call the MPI study "probably normal" in the absence of coronary calcium, but report the perfusion defect in the presence of coronary calcium. Thus, inaccuracy of VECAC in estimating the AS could potentially lead not only to inaccurate characterization of coronary calcium, but also to inaccurate characterization of myocardial perfusion. To assess the effect of VECAC on MPI interpretation in such borderline scans, in 1 center (Columbia), we reassessed MPI in all scans that had been read as having a rest or stress perfusion defect in 1 to 3 segments of the American Heart Association 17-segment model, none of which were described as a “severe” defect. MPI scans were read independently by each reader on 3 separate days, presented in random order on each occasion. On the first day, each reader read the MPI scan, using the 17-segment model. On the second day, each reader was presented with the MPI images, their original interpretation of these images, and the CAC scan images, and asked to reinterpret the MPI images in this context. The third reading day was analogous to the second day, except that CTAC scan images were presented rather than CAC scan images.

**Statistical analysis.** Agreement between VECAC on the CTAC scan and AS, measured on the standard CAC scan and converted to the same 6-level scale, was determined using quadratic weighted kappa statistics, with 95% confidence intervals (CIs) estimated using bootstrap with 1,000 replications. Agreement between readers in VECAC was similarly assessed using quadratic weighted kappa statistics. Vessel-based analysis was performed by evaluating the proportion of segments with correct identification of calcium on the CTAC scan and associated standard kappa scores. Statistical analysis was performed using Stata 10.1 (StataCorp, College Station, Texas).

**Results**

**Agreement between VECAC and Agatston score.** A high degree of association was observed between VECAC and AS, with 63% of VECAC scores in exactly the same category as the AS category and 93% varying by no more than 1 category (Table 1). Only 1 significant outlier occurred, for a patient with marked LAD calcification close to the pulmonary artery, resulting in an AS of 1,082. This was misidentified by 1 reader as pulmonary arterial calcification and thus visually classified as having no coronary calcification. Overall weighted kappa was 0.89 (p < 0.0001). High weighted kappa statistics were observed for each site (Table 2), for women and men (0.89 and 0.90, respectively), and for both low-dose CTAC scans performed on a Philips SPECT/CT scanner (0.90, 95% CI: 0.87 to 0.93) and even lower-dose CTAC scans performed on a Siemens PET/CT scanner (0.88, 95% CI: 0.85 to 0.89).

Statistics for each reader are summarized in Table 2, which reveals minimal variability between readers and sites in measures of agreement. When reviewing the same set of images from Columbia, all 4 readers obtained virtually identical diagnostic performance (Table 3), suggesting no significant intrinsic differences in classificatory ability between readers, which could serve as a confounder for comparisons between sites.

For each AS range, the most common VECAC score was its corresponding score on the 6-level scale, and vice versa, with the exception of AS between 1 and 9. Of the 31 patients with AS 1 to 9, for each of whom VECAC was estimated by 2 readers, VECAC was estimated as 0 in 46 cases, 1 to 9 in 8 cases, and 10 to 99 in 8 cases. Of the 39 cases in which VECAC was estimated as 1 to 9, the actual AS was 0 in 14 cases, 1 to 9 in 8 cases, 10 to 99 in 13 cases, and 100 to 399 (maximum 207) in 4 cases.

**Vessel-based analysis.** Agreement between the CTAC scan and the calcium scoring scan in identifying CAC was good for all vessels. For the left main, LAD, circumflex, and right coronary arteries, percentage agreement was 78.9%.

<table>
<thead>
<tr>
<th>AS</th>
<th>VECAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>323</td>
</tr>
<tr>
<td>1–9</td>
<td>46</td>
</tr>
<tr>
<td>10–99</td>
<td>37</td>
</tr>
<tr>
<td>100–399</td>
<td>5</td>
</tr>
<tr>
<td>400–999</td>
<td>0</td>
</tr>
<tr>
<td>1,000+</td>
<td>1</td>
</tr>
</tbody>
</table>

AS = Agatston score; VECAC = visually estimated calcium artery score.
89.6%, 84.6%, and 83.5%, respectively. Corresponding kappas were 0.41, 0.79, 0.67, and 0.65, respectively. A discrepancy for 1 or both readers in the presence of LAD calcium between CAC and CTAC scans was only observed in 20 of the 197 cases. On subsequent localization analysis, in only 3 of these cases was proximal LAD calcium missed on the CTAC scan and observed on the CAC scan; however in 2 of these, left main calcium had been observed on the CTAC scan by 1 reader. This suggests that VECAC does not systematically miss proximal disease.

**Interobserver reproducibility.** A high degree of interobserver reproducibility in VECAC scores was observed, with readers reporting identical scores in 65% of cases and scores within 1 category of each other in 93% of cases (Table 4). Weighted kappa was 0.90 (95% CI: 0.87 to 0.92, p < 0.0001). High weighted kappas were observed for each site (0.90, 0.88, and 0.90), for women and men (0.89 and 0.90, respectively), and for both low-dose CTAC scans performed on a Philips SPECT/CT scanner (0.90, 95% CI: 0.85 to 0.94) and lower-dose CTAC scans performed on a Siemens PET/CT scanner (0.89, 95% CI: 0.85 to 0.91).

**Effect on interpretation of borderline MPI scans:** borderline MPI imaging was observed in 19 cases. In none of these did use of the CTAC scan rather than the CAC scan result in a difference in MPI interpretation for both readers. In 3 cases, use of CTAC rather than CAC resulted in a difference in MPI interpretation for 1 reader. In 1 case, for Reader 1, an apical defect on stress-only MPI was read as normal when read with the CAC scan, but still read as an apical defect when the CTAC scan was substituted; Reader 2 read this scan as “normal” in all 3 instances. In the second case, Reader 1 called a mild, fixed inferoapical defect that was unchanged with CAC or CTAC images; Reader 2 called the defect “normal” on standard MPI and MPI with CTAC, but called the fixed inferoapical defect when reading MPI with CAC images. In the third case, Reader 1 called a reversible anteroapical defect in the presence of breast attenuation that was unchanged with CAC or CTAC images; Reader 2 called the scan “normal” on standard MPI, “probably normal” on MPI with CTAC, but called the reversible defect on MPI with CAC. Thus, the effect of CTAC rather than CAC images on interpretation of borderline MPI appears to be small, with any discrepancies not consistent between readers.

**Discussion**

**Implications.** The findings of this study have several important implications. First, coronary calcium can be assessed from a low-dose CTAC scan. Although the classification of CAC categories was not identical for the 2 methods, the 93% agreement within 1 category suggests that much of the diagnostic and prognostic information provided by calcium scoring is available from the scan that is already performed.
for attenuation correction in patients undergoing hybrid PET/CT or SPECT/CT MPI. This may in some cases obviate the need for a separate calcium scoring scan, with its attendant modest addition to the total radiation burden of the study.

A second important implication of this study is that it suggests the possibility that standard calcium scoring may be performed using lower radiation dose protocols than have been conventionally used. A recent paper evaluating radiation dose from CAC scoring found that effective doses from standard scanning protocols used in the MESA (Multi-Ethnic Study of Atherosclerosis) and CARDIA (Coronary Artery Risk Development in Young Adults) studies, the International Consortium on Standardization in Cardiac CT (3), and 3 university medical centers, ranged from a low of 0.8 mSv to a high of 10.5 mSv in 1 CARDIA site (4). Effective dose associated with the CTAC scan protocol used in 2 of the 3 sites here was 0.3 mSv, markedly lower than the effective doses used in the corresponding CAC scan protocols (4.7 and 2.9 mSv), despite the greater scan length. Certainly, very-low-dose CAC protocols would require further validation prior to adoption for clinical use. Even so, Mid America Heart has already adjusted its standard clinical protocol to lower radiation dose from calcium scoring. Although cancer risks associated with calcium scoring are extremely low, with estimates for a typical protocol ranging between 0.003% and 0.028% depending upon patient age and sex, screening approximately 50 million individuals with calcium scoring as has been proposed in the SHAPE (Screening for Heart Attack Prevention and Education) guidelines (5) could result in an estimated 5,600 cancers (4), based on the assumptions of the BEIR VII (Biological Effects of Ionizing Radiation VII) report (6). Validation of very-low-dose calcium scoring protocols would require further validation prior to adoption for clinical use. Even so, Mid America Heart has already adjusted its standard clinical protocol to lower radiation dose from calcium scoring.

Benefits of CTAC. Performance of a CTAC scan now potentially adds 3 pieces of diagnostic information to an MPI examination: attenuation correction, extracoronary findings, and CAC classification. Due to the significant amount of attenuation, attenuation correction is regarded as essential for PET MPI, and is routinely employed by all laboratories performing this test. For SPECT MPI, it provides the possibility to eliminate false positive perfusion defects (1), thereby improving test specificity and positive predictive value and avoiding unnecessary downstream testing, including invasive angiography. Reflecting this, a multisocietal position statement maintains that attenuation correction “will improve image quality, interpretive certainty, and diagnostic accuracy” (7). Several studies have now shown the high frequency of relevant extracardiac findings in CT scans performed for cardiac evaluation (8) or attenuation correction (9). Although such findings are incidental in some patients, in others they are causal. Thus, CTAC offers the potential to identify noncardiac causes (e.g., hiatal hernia, aortic dissection, emphysema, and pulmonary edema) of symptomatology, such as chest pain or dyspnea, for which patients present for MPI. As demonstrated here, CTAC also adds the ability to semiquantitatively assess CAC.

Given these 3 additional pieces of diagnostic information resulting from a CTAC scan added to an MPI examination, we propose that the time has arrived for consideration of its reimbursement by the Centers for Medicare and Medicaid Services and third-party payers. Payment generally is deemed reasonable when a procedure incurs extra expense and work, when it is widely used in different settings, and when it has been shown to convey clinical benefit. CTAC scanning improves quality of patient care by: 1) potentially improving diagnostic performance of MPI; 2) decreasing healthcare costs and morbidity from unnecessary follow-up testing; 3) clarifying etiology of presenting symptoms in some patients; and 4) supplementing perfusion imaging with an assessment of the presence and extent of coronary artery disease, which can affect clinical decision making in terms of treatment with antiplatelet and lipid-lowering agents in patients without known flow-limiting coronary disease (10). A complete CTAC study, with documentation of the effect of attenuation correction on perfusion defects, pertinent extracoronary findings, and a VECAC score, merits consideration for an additional billing code.

**CTAC protocol differences between sites.** CTAC scan protocols used in this study differed somewhat between sites. Differences in technique included the use of an end-expiratory breath-hold at Mid America Heart versus free breathing at other sites, and the use of a higher tube current at Columbia. Free breathing can result in gaps in the CT dataset due to movement of the heart cranially with expiration, resulting in missed segments of coronary calcium, whereas lowering tube current results in increased image noise. Nevertheless, there were only minor differences in diagnostic performance between sites, with overlapping 95% CIs for weighted kappa between any pair of sites or readers.

The foremost purpose of the CTAC scan is to provide accurate attenuation correction, and protocol selection must be aimed at ensuring this. Myocardial perfusion images are obtained from an acquisition lasting several minutes, during which the patient is freely breathing, motivating many sites and vendors to prefer a free-breathing CT scan for attenuation correction. Although limited data suggest that an
end-expiratory CT scan may be as effective for attenuation correction as a free-breathing scan (11,12), this issue requires further study, and the optimal breathing protocol for CTAC remains uncertain (13). Even if end-expiratory CT scanning is preferred, in practice, many patients are unable to effectively follow instructions to reliably perform an end-expiratory scan.

It also remains unclear what tube current is necessary to ensure optimal CTAC. Columbia, using the vendor’s suggested protocol, typically employed effective mAs of 50, whereas the other sites used ~10 mAs, resulting in appreciably noisier images (Fig. 2). Nevertheless, the experience is that these noisier images are generally adequate for PET attenuation correction. The spatial resolution of the CT scan is much higher than that of Rb-82 PET or Tc-99m SPECT, resulting in image noise being decreased by averaging CT numbers in the larger voxels used for nuclear imaging. PET/CT data obtained from N-13 ammonia assessment of myocardial blood flow evaluation (14) and F-18 fluorodeoxyglucose metabolic imaging (15), as well as phantom work (16), suggest that very-low-dose CTAC is indeed adequate; however, further validation for myocardial perfusion would be desirable.

Study limitations. The 6-level system used typically resulted in accurate classifications, with the exception of patients classified with VECAC of 1 to 9, who usually had AS of either 0 or 10 to 99, with similar representation in each group (14 vs. 13 cases). Although a grading system with fewer categories could be considered, by grouping a VECAC of 1 to 9 together with either VECAC of 0 or VECAC of 10 to 99, each of these approaches would have limitations in terms of risk stratification (17).

A lower kappa value was noted for the left main than for other coronary arteries. The left main can be difficult to distinguish from the proximal LAD or circumflex, especially on low-dose and non-breath-hold scans.

Although we observed generally good inter-reader reproducibility, all 6 readers were highly experienced in reading standard CAC scores prior to initiation of the current study. The subjective assessment involved in assigning a VECAC score offers the potential for more variability among less experienced readers. A possible solution to this limitation would be a true quantitative measure of CAC on the CTAC scans, obtained by marking regions of potential interest identified with coronary calcification and performing some calculation based on these regions. Use of a CT number threshold of 130 Hounsfield units, as in the AS (18), would result in underidentification of some areas of coronary calcification. In our experience, due to CTAC scan technique, and especially free breathing, which may blur coronary arteries and thereby result in decreased CT numbers, some regions with mean CT numbers of even 80 Hounsfield units appear consistent with CAC. This differs from the findings of Wu et al. (19), who observed that a standard Agatston method could be reliably used on 0.9 mSv CT scans performed for lung cancer screening. For CTAC scans, any threshold chosen and calculation performed would likely need to vary depending on scan technique. Although this is a potentially important area for future study, our results suggest that the semiquantitative assessment provided by VECAC scoring provides outstanding agreement with AS while generally requiring little time for assessment.

Another limitation related to reproducibility is that we did not assess interscan reproducibility for VECAC or AS. Since patients were already obtaining 2 CT scans of the chest, 1 for CTAC and 1 for AS, we did not consider it justifiable in terms of radiation burden to perform 2 additional scans to assess interscan reproducibility. Nevertheless, it is known that repetition of calcium scoring is associated with a small difference in AS between scans. With a 16-slice scanner, Horiguchi et al. (20) observed a mean interscan difference in AS of 7% with low motion artifact and 19% with high motion artifact, whereas with a 64-slice scanner, Matsuura et al. (21) observed a mean interscan difference in diastole of 8%. In utilizing a 6-level VECAC score rather than a precise quantification of calcium, our goal was to stratify patients into broad categories corresponding to levels of cardiovascular risk, rather than to precisely estimate AS. Nevertheless, the effect of interscan variability of AS on the estimation of AS by VECAC requires further study.

Another potential limitation relates to the generalizability of our findings. Each site in our study used a hybrid scanner with at least 16 slices for CT. Systems with fewer slices may be more susceptible to motion artifacts, especially with free breathing (22). Although standard multidetector row scanners with ≥4 slices and gantry rotation time ≤0.5 s are generally regarded as adequate for calcium scoring, the agreement of VECAC with AS for scanners with <16 slices has not been studied. Some SPECT/CT scanners only allow tube currents considerably lower than those used in this study (e.g., maximum 2.5 mA), increased tube voltage, slow rotation speed, and decreased spatial resolution (23). This combination of scan parameters will likely result in decreased accuracy of VECAC. Thus, estimation of VECAC from CTAC images on such scanners should not be performed in the absence of independent validation of its diagnostic performance.

Conclusions

In summary, CAC can be visually estimated from a CTAC scan with outstanding agreement with AS and a high degree of interobserver reproducibility. CTAC scanning should be routinely used when available, and the associated images inspected. This approach supplements the diagnostic information from MPI by correcting perfusion defects caused by soft-tissue attenuation, demonstrating extracoronary causes of chest pain and dyspnea, and characterizing CAC.
REFERENCES


Key Words: coronary calcium • hybrid imaging • PET • SPECT.