UCLA
UCLA Previously Published Works

Title

Estimating the accuracy of dual energy chest radiography for coronary calcium detection with lateral or anteroposterior orientations

Permalink

https://escholarship.org/uc/item/5sg546w7

Journal

Medical Physics, 49(9)

ISSN

0094-2405

Authors

Hsieh, Scott S
Budoff, Matthew J

Publication Date

2022-09-01

DOI

10.1002/mp.15855

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Estimating the accuracy of dual energy chest radiography for coronary calcium detection with lateral or anteroposterior orientations

Scott S. Hsieh[1]*, Matthew J. Budoff[2]
[1]Department of Radiology, Mayo Clinic Rochester,
[2]Department of Medicine, Lundquist Institute at Harbor-UCLA,
* Corresponding author: hsieh.scott@mayo.edu

Purpose. Coronary artery calcium (CAC) scoring with CT has been studied as a risk stratification tool for cardiovascular disease. However, concerns remain from the radiation dose, economic expense, and incidental findings associated with this exam. Dual energy chest x-ray (DE CXR) has been proposed as an alternative, but validation of this technique remains limited. The purpose of this work was twofold: first, to estimate the sensitivity and specificity of DE CXR using simulation of patient datasets in a CAC screening cohort; second, to assess if sensitivity and specificity could be improved using a lateral instead of an anteroposterior (AP) orientation.

Methods. Starting from a cohort of 200 CAC scoring CT exams, we excluded patients with metal wires, data truncation, or with age outside 40-75 years. The first half were used for reader training, while the second half were reserved for blinded validation. After exclusions, 73 patients remained in the validation set. The fraction of CT CAC scores in the validation set of 0, 1-99, 100-299, and 300+ were 36%, 25%, 14%, and 26%, respectively. CT datasets were decomposed on a voxel-by-voxel basis into mixtures of water and calcium according to CT number. DE CXR images were simulated using polyenergetic forward projection with scatter estimated from Monte Carlo. We assumed a technique of 60 and 120 kVp for the dual energy acquisition. The tube current was scaled such that the estimated radiation dose from DE CXR was 10 times less than CAC scoring CT. Patient motion was not simulated. Two readers read the validation set in a blinded, randomized fashion, and estimated the amount of CAC in each DE CXR image using a semiquantitative 4-point scale. Although patients present on a spectrum of CAC severity, in the primary analysis, sensitivity and specificity were calculated by dichotomizing patients into two categories of CT CAC (Agatston) scores of either 0-99 or 100+.

Results: From the lateral orientation, average sensitivity between two readers was 69% (range, 69%-69%), specificity was 85% (range, 84%-86%), and area under the curve (AUC) was 0.81 (range, 0.80-0.81). From the AP orientation, average sensitivity was 35% (range, 31%-38%), average specificity was 70% (range, 66%-73%), and AUC was 0.54 (range, 0.53-0.55). Reader DE CXR scores agreed within 1 point of the 4-point scale on 97% of ratings from the lateral orientation and 80% from the AP orientation. From the lateral orientation, AUC increased when considering higher CT CAC score thresholds as disease positive; for thresholds of 1+, 300+, and 1000+, average AUC was 0.72, 0.81, and 0.92 respectively. From the AP orientation, AUC was 0.57, 0.55, and 0.61, respectively.

Conclusions: DE CXR for CAC scoring may have higher diagnostic accuracy when acquired from the lateral orientation. The sensitivity and specificity of lateral DE CXR, when combined with its modest cost and radiation dose, suggest a possible role for this technique in screening coronary calcium in lower risk individuals. These estimates of diagnostic accuracy are derived from simulation of patient datasets and have not been corroborated with experimental or clinical images.

I. INTRODUCTION

Cardiovascular disease (CVD) has remained the leading cause of death in the United States for nearly a century. While several interventions exist to reduce the risk of CVD, most interventions are associated with costs and possible adverse events. A guiding principle in the management of CVD is that interventions should be tailored to individual risk (1). For patients at very high risk of CVD, surgical options may be appropriate. For patients at moderate risk of CVD, a wide variety of preventative options can be used: statins can be prescribed (2) or lifestyle interventions such as diet and exercise can be pursued (3). The paradigm of personalizing the intervention to individual risk requires accurate predictions of risk. Models such as the Framingham risk score (4) provide risk estimates from variables such as age, gender, smoking status, cholesterol levels, and blood pressure, but there remains unexplained variability in cardiovascular events.

Coronary artery calcium (CAC) is an indicator for atherosclerosis (5, 6) and has shown to provide the most additional information among emerging CVD biomarkers. CAC scoring in CT is usually performed using the original system from Agatston, although other scoring systems have also been described (7). One study showed that CAC measurement by CT provided four times better net reclassification improvement than high-sensitivity C-reactive protein (hs-CRP), carotid media-intima thickness, or family history (8).

Nonetheless, the use of CAC scoring has been controversial (1). Several minor harms have been attributed to CAC scoring CT. The associated radiation dose of approximately 1 mSv is equal to several months of background radiation. Incidental findings are common and may generate anxiety and distress due to overdiagnosis and overtreatment. Finally, the economic cost of CAC scoring is greater than currently used tools including cholesterol and blood pressure

measurement. Because of these minor harms, CAC scoring has not been universally endorsed but has only been recommended under select conditions (9).

Dual energy chest x-ray (DE CXR) has been proposed as an alternative form of CAC scoring that minimizes these harms. DE CXR is attractive from a translational standpoint because some CXR systems are already equipped with dual energy and has the potential to be widely available if recommended. However, studies on the performance of DE CXR for CAC screening have been limited.

Xu et al. explored detection of CAC using DE fluoroscopy in an experimental study on phantoms. They showed detectability of relatively small calcifications (<100 CT CAC score) even at very low doses (10). Zhou et al. used numerical simulations to estimate artifacts using DE CXR, finding that errors introduced by beam hardening or scatter were modest and could be controlled (11). These studies demonstrate that in ideal laboratory conditions, DE CXR is able to detect CAC, but they do not prove that DE CXR would be effective in human patients who may present with extracoronary calcium that may complicate the detection of CAC.

Mafi et al. retrospectively evaluated DE CXR acquired on a population of 39 patients, creating a quantitative score that was well correlated with CT CAC score for patients with high calcium (>400 CT CAC score) (12). However, there was no correlation at lower calcium levels. Song et al. evaluated DE CXR on a larger cohort of patients together with specialized software to improve the contrast of CAC and compensate for patient motion (13). They found that DE CXR could generally differentiate patients with severe CAC (>300 CT CAC score) from minimal CAC (<50 CT CAC score). However, their patient population was also enriched with severe disease due to the nature of their retrospective study.

The purpose of this work is to estimate the accuracy of DE CXR for the detection of CAC in a more typical screening cohort using simulations from CT source data. We further hypothesized that because CAC must be differentiated from other extracoronary calcifications, a

lateral (LAT) acquisition might outperform an anteroposterior (AP) acquisition by reducing the spatial overlap of CAC from other types of calcium.

II. METHODS

II. A. Patient cohort

A series of 200 CT exams were collected from the Lundquist Institute from February 8 to March 16, 2021. The first 100 exams were designated for reader training, and the second 100 exams were designated for blinded validation.

The exams were screened for four exclusion criteria: (1) metallic wires, (2) arms down at the side, (3) field of view truncation, and (4) ages less than 40 years or greater than 75 years. The rationale for these exclusion criteria were as follows. Metal wires, present in some undergarments, create strong and distracting signals in dual energy processing. However, these wires could be excluded in clinical deployment of DE CXR with appropriate patient preparation. The presence of arms down at the side increases patient comfort for CT but creates additional attenuation in the lateral acquisition. In clinical deployment, the arms could be elevated with appropriate patient instructions and with arm support in some cases. Field of view truncation limits simulation accuracy for DE CXR as the dataset is incomplete. In screening the cases, small amounts of truncation of a few centimeters of tissue were allowed; more severe truncation led to exclusion. Finally, patients with ages outside the range of 40-75 years are not ordinarily recommended for CAC screening, and we felt that they were not representative of the target application of DE CXR. Using these criteria, a total of 27 patients were excluded in the reader validation set. Excluded datasets were not further processed into simulated DE CXR images. After exclusions, the average age of the validation cohort was 59 years (standard deviation, 9 years), and 46 of 73 patients were male.

All CT exams were acquired with a 256-slice Revolution CT in a single axial scan. While Agatston scoring was performed using a conventional 3 mm slice thickness in a smaller field of view, our DE CXR simulations used a slice thickness of 0.625 mm and a 50 cm field of view. The Lunquist Institute is a high volume CAC scoring center and the CT protocols were dose optimized, with the coronary arteries positioned near isocenter. Scans were performed at 120 kVp at a mean of 66 mAs (standard deviation, 32 mAs).

Figure 1 shows the distribution of CT CAC score for the validation cohort. The median CT CAC score was 23. Past research has established changes in management for patients with a CT CAC score of 0 as a negative risk factor (14), or 100+ as a positive risk factor (9). We prespecified these thresholds as clinically important, and we prespecified that the primary endpoint was the sensitivity and specificity for these thresholds. Scores of 1000+ have also been associated with very high risk (15) but have low prevalence in the general population. The number of patients in validation cohort with CT CAC scores of 0, 1-100, 101-300, and 301+ were 26 (36%), 18 (25%), 10 (14%), and 19 (26%), respectively. For comparison, the corresponding distribution of CT CAC scores in the Multi-Ethnic Study of Atherosclerosis (MESA) was 51%, 26%, 11%, and 12%, respectively (6). Our cohort had fewer patients with 0 score and more patients with 301+ scores compared to MESA, which could be attributed to the patient recruitment process, as self-referred patients are scanned by the Lundquist Institute who may have high self-perceived CVD risk. Nonetheless, the distribution of CAC scores in our cohort is more representative of a screening population than that of past retrospective clinical studies (12, 13).

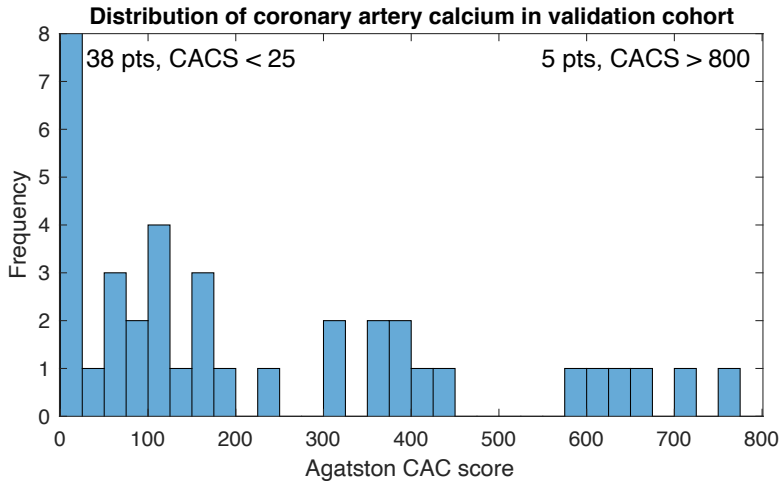

Figure 1. Histogram of coronary artery calcium (CAC) scores as measured by CT. Bin width is 25. 38 patients are present in the first bin, but to improve visibility for the remainder of the histogram, the bar height is truncated. 5 patients presented with scores above 800 (range, 1300-3200) and are not shown on this histogram.

II. B. Dual energy chest x-ray (DE CXR)

From the CT datasets, we simulated dual energy chest x-ray (DE CXR) images. Figure 2 shows several steps in our simulation process.

The CT dataset was first preprocessed. The bottom and top 15 slices were eliminated because these slices were incomplete and partially masked due to the nature of the axial scan. The volume was then padded by the outermost slice to attain a total of 256 slices. The patient table and calcium calibration rods below the patient were deleted as they would not be present in a CXR acquisition. To do this, a mask containing the largest single connected component of the volume greater than -500 HU was selected to eliminate the rods. This mask was processed by image opening using a disk filter with radius of 3 pixels to eliminate the table, and a hole filling operation was then performed to include the lungs. Data outside the mask was set to air. Figure 2b shows an example of this mask process.

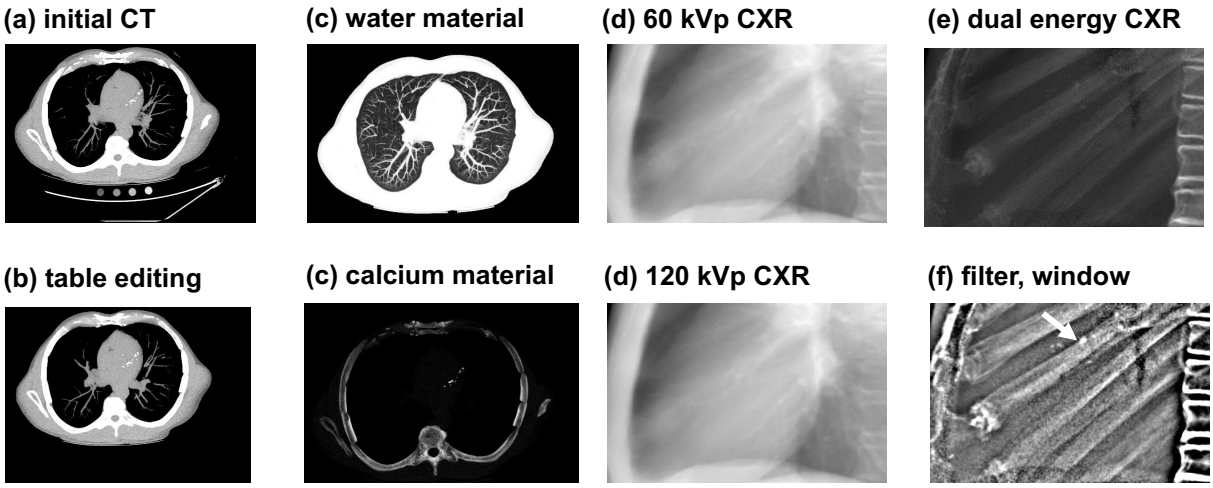

Figure 2. Major steps of our image processing pipeline to simulate dual energy chest x-ray images. (a) Representative maximum intensity projection through 30 slices of the initial CT scan. (b) After table and calcium contrast vial elimination. (c) (Top) Water and (bottom) calcium material decomposition, determined on a voxel-by-voxel basis. (d) Simulated, polychromatic chest x-ray images from forward projection of water and calcium material volumes. Scatter has been added through a Monte Carlo program and attenuated by a scaling factor from an assumed anti-scatter grid. (e) Dual energy chest x-ray produced by a linear combination of the 60 and 120 kVp chest x-rays to highlight calcium signal. (f) Bandpass filtering of the dual energy chest x-ray, with low frequencies eliminated to subtract the background and high frequencies suppressed to reduce noise. The final image is windowed narrowly to highlight coronary calcium. Arrowhead points to coronary calcium.

To perform dual energy simulations, it was necessary to perform basis material decomposition on the original single-energy CT dataset. This decomposition was performed on a voxel-by-voxel basis, and with the assumption that voxels less than 100 HU were composed entirely of water of the appropriate density, and that excess attenuation above 100 HU was due entirely to calcium, with a density estimated using an effective energy of 60 keV. Specifically, for a voxel with CT number x, we chose the density of water in that voxel to be $\min\left(1.1, 1+\frac{x}{1000}\right)$ g/cm^3, and the density of calcium to be $\max\left(0, \frac{x-100}{3195}\right)$ g/cm^3. The normalization factor for calcium was chosen because calcium is 3.195 times more attenuating than water at 60 keV (16). Figure 2c shows an example of the water and calcium basis materials.

After decomposition into water and calcium, forward projection was performed using geometric ray tracing. A point source and pixelated detector was used. The total attenuation for

each ray was estimated by stepping along the ray connecting the source to a detector pixel and applying bilinear interpolation at each step. We assumed a geometry of DE CXR with a source-detector distance of 180 cm and a source-isocenter distance of 150 cm. The detector was 22.5 cm wide and 14.4 cm high, with pixel size of 0.5 mm. Real CXR systems may have higher resolution that could potentially enhance the boundaries of small CAC deposits. However, our ability to simulate fine features was limited by the resolution of the underlying CT dataset.

We assumed that the DE CXR system acquired images at 60 kVp and 120 kVp, matching the protocol of an existing DE CXR system. The x-ray spectrum was estimated using the Spektr v3 toolkit (24). The mAs used was proportional to the reference mAs R from the CT scan at 120 kVp. Specifically, we used $2R$ at 60 kVp and $R/3$ at 120 kVp. For the average CT scan at 66 mAs, the corresponding number of photons simulated was 6.2 * 10^6 per mm^2 at the detector at 60 kVp and 7.6 * 10^6 per mm^2 at the detector at 120 kVp.

The dose of DE CXR compared to CT can be estimated as follows. First, the source-isocenter distance of the Revolution CT scanner is 2.4 times smaller than that of our simulated DE CXR (62.5 vs 150 cm), implying a 5.75 reduction in dose using the inverse square law. Second, the total energy of the beam at 60 kVp (assuming all photons are completely absorbed) is 10.9 times less than that of 120 kVp as estimated from the Spektr spectra. When these factors are combined with the mAs used, we calculate an estimate dose for DE CXR that is 11 times smaller than that of the original CAC scoring CT.

Scatter was estimated using MCGPU, an open source Monte Carlo package (17). To reduce the computational runtime, volumes were downsampled by a factor of 4 and the code was executed on a remote computer cluster. The volume was also padded using the top and bottom slice to attain a total volume thickness of 32 cm to model photons that first scatter within the imaging volume, then scatter outside the imaging volume, and finally are received by the detector. A total of 10^{10} photons were simulated using Monte Carlo for each kVp (60 or 120 kVp) and

orientation (lateral or AP). To reduce the noise from the Monte Carlo scatter simulation process, the estimated scatter was blurred by a Gaussian with standard deviation of 3 pixels.

Scatter complicates dual energy processing, because the amount of detected scatter depends on patient anatomy, may be different between lateral and AP acquisitions, and may differentially affect low and high energy images. Although software scatter corrections could mitigate bias resulting from scatter, implementation of such a correction is nontrivial. We instead simulated an anti-scatter grid that would eliminate most of the scatter and hence simplify the dual energy processing. The anti-scatter grid had a selectivity factor of 11.9 and a primary transmission factor of 71%. The transmitted primary and scatter signals were scaled accordingly prior to noise addition. Noise was added to the primary photons in a compound Poisson process at different energy levels using a 3 keV energy step size. Although the scatter photons are polyenergetic, the energy distribution of the scatter was not easily exported from MCGPU. For the purpose of Poisson noise addition, we assumed that all scatter photons had an energy of 48 keV for the 60 kVp acquisition and 76 keV for the 120 kVp acquisition. Figure 2d shows examples of the simulated CXR images after including scatter.

Both the 60 kVp and 120 kVp acquisition were normalized to air and log-transformed, yielding images we denote as Im_{60} and Im_{120}. The calcium image was a linear combination of these images, $Im_{60} - 1.25\, Im_{120}$, where the factor of 1.25 was chosen to cancel water in a 30 cm water object. Figure 2e shows an example of this linear combination. In order to increase the conspicuity for small calcifications and improve the dynamic range of the image shown to the readers, we performed background subtraction by subtracting a blurred version of the calcium image, using a Gaussian of standard deviation 5 mm, from the original image. Finally, we performed noise reduction by blurring this image by 0.5 mm. Figure 2f shows an example of the bandpass filtered images with windowing applied to highlight CAC. These image postprocessing steps were determined using the training cohort only and were fixed prior to simulating the blinded validation cohort.

II. C. Reader training and scoring

Two readers evaluated the validation images. Reader 1 was a cardiologist (M.J.B.) with more than 20 years of experience interpreting clinical CT scans for CAC scoring. Reader 2 (S.S.H.) was a physicist who developed the concept and performed the simulations.

Readers were shown the training dataset first with ground truth CT CAC scores provided. Some of these images had the CT CAC scores withheld, and the readers practiced reading these images first before seeing the answers and discussing how they arrived at their scores. After reviewing the training images, readers both agreed on a 4-point scale for the validation dataset. The 4 points on the scale corresponded to CAC scores of 0, 50, 200, and 500, which hereafter will be referred to as DE CXR CAC scores. The points on this scale correspond to CT CAC scores of 0, 1-99, 100-299, and 300+, but because readers did not provide more granular numerical scores, we use only the labels 0, 50, 200, and 500 in this work to refer to DE CXR CAC scores.

The validation cohort was shown in a blinded fashion, with the AP and LAT DE CXR images shown separately and independently scrambled. Readers could not correlate an AP view with a LAT view but independently scored each. The training materials were available as a reference during the reading session.

Reader 2 also rated the noise of each image subjectively using a 1 to 5 Likert scale, under the hypothesis that the noise in some DE CXR images would render the reader scores unreliable. This noise rating was performed prospectively, prior to unblinding and data analysis.

III. Results

Figure 3 shows five examples of simulated DE CXR from the validation set, along with reader scores and ground truth CT CAC scores. Examples were selected across the spectrum

of CAC severity and included cases with false negatives (undercall), false positives (overcall), and correct detection. The far right case is an example where the dose of lateral DE CXR may have been too low to provide accurate CAC detection. In general, CAC was more difficult to differentiate from other calcifications from the AP orientation.

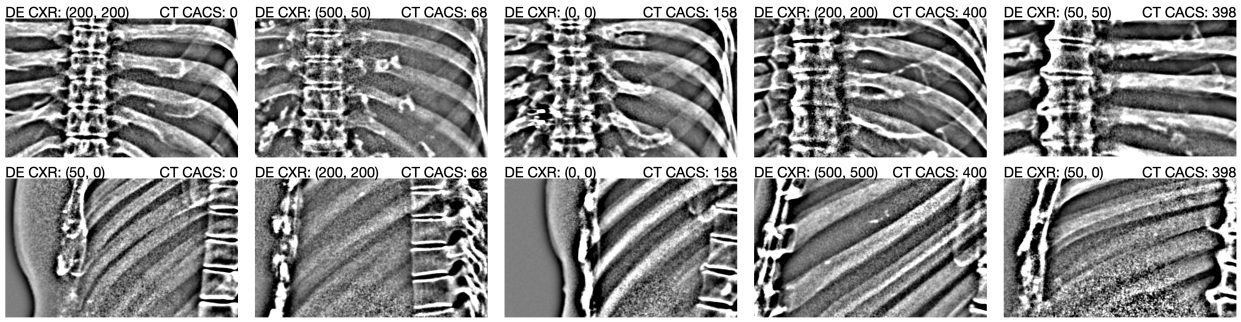

Figure 3. Example cases shown to readers. Each column is one case. Top row is lateral DE CXR, bottom row is AP DE CXR. Reference standard CT CAC score (CT CACS) is shown at the top right of each image. Reader 1 and reader 2 scores are shown, in order, at the top left of each image. (Far left) A case with no CAC. Both readers overcalled this case in AP DE CXR, but provided scores of 0 or 50 in lateral DE CXR. (Center left) A case with a small amount of CAC. Reader disagreed in AP, with one reader overcalling CAC. Calcification is visible from the lateral view. (Center) A case with moderate CAC that was missed by both readers in both AP and lateral. (Center right) A case with severe CAC, easily seen in lateral but with more ambiguity in AP. (Far right) A case with severe CAC and a high noise rating of 4 for the lateral image. 8 of 73 lateral images had noise ratings of 4 or greater. Noise may have led to a false negative.

Table 1 and Table 2 compare reader scores for DE CXR to reference standard CT CAC scores from lateral and AP orientations, respectively. From these tables we may compute the sensitivity and specificity of DE CXR for detecting CT CAC scores of 100+, which is the primary endpoint of this study. For the lateral orientation, the sensitivity of DE CXR was 69% for both readers, and the average specificity was 85% (84% for Reader 1, 86% for Reader 2). For the AP orientation, the average sensitivity was 35% (38% for Reader 1, 31% for Reader 2), while the average specificity was 70% (66% for Reader 1, 73% for Reader 2). Both the sensitivity and the specificity were improved by the lateral orientation.

Figure 4 shows receiver operating characteristic (ROC) curves for both readers and for both orientations. The area under the curve (AUC) for the lateral orientation and for Reader 1

and Reader 2 for the lateral orientation was 0.80 and 0.81, respectively. For the AP orientation, it was 0.53 and 0.55, respectively, indicating performance that was only slightly better than chance.

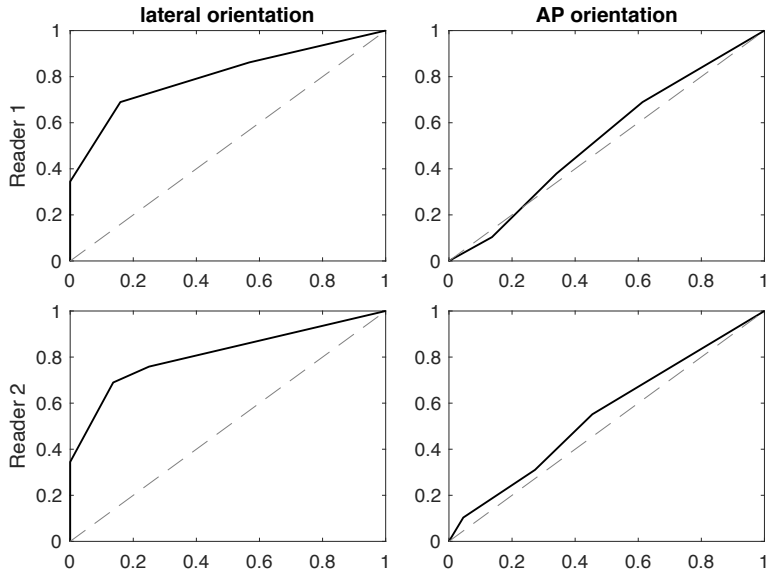

Figure 4. Receiver operating characteristic (ROC) curves for the task of detecting cases with CT CAC scores of 100+. Rows correspond to readers and columns correspond to orientations. For each subplot, the y-axis is sensitivity and x-axis is 1 – specificity, and the gray dashed line corresponds to the performance of random guessing.

Table 1. Contingency table comparing reader CAC estimates from DE CXR acquired from the lateral orientation to that from original CT. All values are presented as averages (Reader 1, Reader 2).

Lateral	CT, 0	CT, 1-100	CT, 101-300	CT, 301+	Sum
DE CXR, 0	16 (9, 23)	10 (10, 10)	2.5 (2, 3)	3 (2, 4)	31.5 (23, 40)
DE CXR, 50	7.5 (13, 2)	4 (5, 3)	2 (2, 2)	1.5 (3, 0)	15 (23, 7)
DE CXR, 200	2.5 (4, 1)	4 (3, 5)	4.5 (5, 4)	5.5 (5, 6)	16.5 (17, 16)
DE CXR, 500	0 (0, 0)	0 (0, 0)	1 (1, 1)	9 (9, 9)	10 (10, 10)
Sum	26	18	10	19	73

Table 2. Contingency table comparing reader CAC estimates from DE CXR acquired from the AP orientation to that from original CT. All values are presented as averages (Reader 1, Reader 2).

AP	CT, 0	CT, 1-100	CT, 101-300	CT, 301+	Sum
DE CXR, 0	13 (10, 16)	7.5 (7, 8)	4 (4, 4)	7 (5, 9)	31.5 (26, 37)
DE CXR, 50	6.5 (8, 5)	3.5 (4, 3)	2.5 (3, 2)	5.5 (6, 5)	18 (21, 15)
DE CXR, 200	5 (5, 5)	4.5 (4, 5)	3.5 (3, 4)	3.5 (5, 2)	16.5 (17, 16)
DE CXR, 500	1.5 (3, 0)	2.5 (3, 2)	0 (0, 0)	3 (3, 3)	7 (9, 5)
Sum	26	18	10	19	73

Table 3. Three measures of diagnostic accuracy for detection of different CT CAC thresholds from a lateral acquisition. AUC of ROC is the area under the receiver operating characteristic curve. All values presented as average value (reader 1 value, reader 2 value). For each threshold, patients are dichotomized into disease present or absent according to the CT CAC threshold. The second column (N_+/N_{tot}) indicates the number of disease positive patients compared to the total population of 73 patients. Sensitivity and specificity are calculated using the nominal reader labels of 0, 50, 200, and 500. Because there was no label for 1000+, the sensitivity and specificity for this threshold is not reported.

Threshold	N_+/N_{tot}	Sensitivity	Specificity	AUC of ROC
1+ CAC	47/73	67% (70%, 64%)	62% (35%, 88%)	0.72 (0.65, 0.78)
100+ CAC	29/73	69% (69%, 69%)	85% (84%, 86%)	0.81 (0.80, 0.81)
300+ CAC	19/73	47% (47%, 47%)	98% (98%, 98%)	0.81 (0.81, 0.82)
400+ CAC	13/73	62% (62%, 62%)	97% (97%, 97%)	0.90 (0.91, 0.90)
1000+ CAC	5/73	N/A	N/A	0.92 (0.92, 0.92)

Table 3 shows three measures of diagnostic accuracy (sensitivity, specificity, and AUC) for other CT CAC score thresholds from the lateral orientation. For brevity, the corresponding sensitivity and specificity for the AP orientation are not reported here, but these metrics could be calculated from Table 2. For the AP orientation, the AUC averaged between both readers was 0.57, 0.54, 0.55, 0.56, and 0.61 for CT CAC scores thresholds of 1+, 100+, 300+, 400+, and 1000+. For every case considered, all three measures of diagnostic accuracy were higher from the lateral orientation than from the AP orientation.

For the lateral orientation, AUC improved with higher CAC thresholds, increasing from 0.72 for the task of detecting any (1+) CAC to 0.92 for the task of detecting very severe (1000+) CAC, although very few patients presented with very severe CAC. Table 3 also reports diagnostic accuracy metrics for a CT CAC threshold of 400+ instead of 300+. Some previous studies on CT CAC have used a threshold of 400 to differentiate moderate from severe CAC (18). The improvement seen between 300+ and 400+ in this study may be a consequence of the distribution of CAC within this cohort as seen in Figure 1.

For the three summary metrics of diagnostic accuracy, inter-reader agreement was excellent in most cases examined. For the individual scores provided to DE CXR images, inter-reader agreement was more variable. Of the 73 lateral DE CXR images, both readers agreed exactly on the 4-point scale for 59% of images and within one point for 97%. Of the 73 AP DE

CXR images, readers agreed exactly on 48% and within one point on 80%. The most common disagreement was in the use of the 0 or 50 score: Reader 1 used a score of 50 more frequently than did Reader 2 in cases where the presence of CAC was ambiguous. This may explain the differences in sensitivity and specificity seen in Table 3 for the task of detecting 1+ CAC.

Table 4 shows the impact of censoring lateral images with high noise, as subjectively rated using the Likert scale. When the 8 patients with the highest noise ratings were censored, the sensitivity, specificity, and AUC increased incrementally. Further improvements in AUC were seen when the 6 patients in the next highest noise ratings were censored. An example of a noise rating of 4 is shown in Figure 3. It is possible that for these censored patients, better CAC detection could have been achieved with greater mAs. Images from the AP orientation were rarely rated as high noise due to the reduced tissue path-lengths, and because the mAs from the AP orientation was equivalent to that of the lateral orientation.

Table 4. Diagnostic accuracy measures for detecting patients with 100+ CT CAC scores from the lateral orientation after excluding patients with high noise ratings.

Noise rating	N_+/N_{tot}	Sensitivity	Specificity	AUC of ROC
Any	29/73	69% (69%, 69%)	85% (84%, 86%)	0.81 (0.80, 0.81)
1 to 3 only	25/65	74% (72%, 76%)	86% (85%, 88%)	0.82 (0.81, 0.82)
1 or 2	23/59	76% (74%, 78%)	85% (83%, 86%)	0.83 (0.81, 0.84)

Discussion

In this reader study evaluating simulated images, DE CXR for CAC detection had higher diagnostic accuracy when performed from the lateral orientation than from the AP orientation. For the task of detecting patients with CT CAC scores of 100+, sensitivity from the lateral orientation (69%) was nearly twice that of AP (35%), and specificity was improved (85% compared to 70%), and AUC was improved (0.81 compared to 0.54). Agreement between both readers was excellent for these measures of diagnostic accuracy. Compared to prior work (10-13), our study

demonstrates the potential value of DE CXR in a screening cohort with CAC prevalence that is more representative of the general population.

Due to the differing costs and benefits, DE CXR for CAC scoring may have a different role than CT. Patients at intermediate risk for cardiovascular disease (CVD) are currently recommended to undergo CT because CT can both identify patients at high risk for treatment intensification as well as patients at low risk for treatment de-escalation. DE CXR may have limited utility in this cohort because it cannot reliably differentiate patients with a CAC score of 0 from a score of 1-99. On the other hand, CAC scoring by CT is not routinely recommended for healthy patients at low risk of CVD, even though one study found that 70% of young adults hospitalized for coronary heart disease would have been classified as “low risk” (19). For this low-to-intermediate risk cohort, DE CXR could be applied as a first screening tool, as a CAC score of 0 would not change management, but a score of 100+ could trigger additional intervention. Because of the imperfect sensitivity and specificity of DE CXR, it may be necessary to follow a positive exam from DE CXR with a conventional CAC scoring CT.

Major limitations stemming from the simulation nature of this study include the accuracy of decomposing single-energy CT images into water and calcium basis materials and the absence of patient motion. While we believe we have made reasonable assumptions in estimating calcium content, it is not mathematically possible to determine calcium content from single energy images. Patient motion is expected to lead to misregistration between low and high energy images. Sliding organ techniques could reduce misregistration (13). Gating and faster switching between low and high energies could also reduce misregistration. In cone beam CT, kVp switching from 81 to 125 kVp has been implemented at 30 fps on hardware not specifically designed for kVp switching (20), suggesting that certain CXR systems with digital detectors could be transformed into dual energy platforms with only a software upgrade. In diagnostic CT, switching times between 80 and 140 kVp have been reduced to under a millisecond in some systems, which could nearly eliminate

misregistration. Spectral detectors could completely eliminate misregistration, although the availability of these detectors is currently limited (21).

To overcome the limitations of simulation, it may be possible to acquire experimental spectral images from new photon counting CT scanners. These scanners acquire low-dose scout scans (sometimes called topograms) for patient positioning, and spectral topograms have already been evaluated for other applications (22). Spectral topograms may be able to detect CAC in the context of a gated chest photon counting CT exam, where reference standard CAC scores will also be immediately available, although the low dose of topograms may present a challenge. Spectral topograms may not have the dose or motion characteristics of DE CXR but could provide additional experimental evidence and motivate larger clinical studies to directly test the efficacy of DE CXR for screening of coronary calcium in lower risk individuals.

Acknowledgments. We acknowledge support from the National Institutes of Health (R21 EB026790).

Data availability statement. The reader scores and simulated DE CXR images are available upon request.

References

1. Weintraub WS, Diamond GA. Predicting cardiovascular events with coronary calcium scoring. Mass Medical Soc; 2008. p. 1394-6.
2. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Journal of the American College of Cardiology. 2014;63(25 Part B):2889-934.
3. Lin JS, O'Connor E, Evans CV, Senger CA, Rowland MG, Groom HC. Behavioral counseling to promote a healthy lifestyle in persons with cardiovascular risk factors: a systematic review for the US Preventive Services Task Force. Annals of internal medicine. 2014;161(8):568-78.
4. Lloyd-Jones DM, Wilson PW, Larson MG, Beiser A, Leip EP, D'Agostino RB, et al. Framingham risk score and prediction of lifetime risk for coronary heart disease. The American journal of cardiology. 2004;94(1):20-4.
5. Budoff MJ, Georgiou D, Brody A, Agatston AS, Kennedy J, Wolfkiel C, et al. Ultrafast computed tomography as a diagnostic modality in the detection of coronary artery disease: a multicenter study. Circulation. 1996;93(5):898-904.
6. Detrano R, Guerci AD, Carr JJ, Bild DE, Burke G, Folsom AR, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. New England Journal of Medicine. 2008;358(13):1336-45.
7. Criqui MH, Denenberg JO, Ix JH, McClelland RL, Wassel CL, Rifkin DE, et al. Calcium density of coronary artery plaque and risk of incident cardiovascular events. Jama. 2014;311(3):271-8.
8. Yeboah J, McClelland RL, Polonsky TS, Burke GL, Sibley CT, O'Leary D, et al. Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals. Jama. 2012;308(8):788-95.
9. Lloyd-Jones DM, Braun LT, Ndumele CE, Smith Jr SC, Sperling LS, Virani SS, et al. Use of risk assessment tools to guide decision-making in the primary prevention of atherosclerotic cardiovascular disease: a special report from the American Heart Association and American College of Cardiology. Circulation. 2019;139(25):e1162-e77.
10. Xu T, Ducote JL, Wong JT, Molloi S. Feasibility of real time dual-energy imaging based on a flat panel detector for coronary artery calcium quantification. Medical physics. 2006;33(6Part1):1612-22.
11. Zhou B, Wen D, Nye K, Gilkeson RC, Eck B, Jordan D, et al. Detection and quantification of coronary calcium from dual energy chest x-rays: Phantom feasibility study. Medical physics. 2017;44(10):5106-19.
12. Mafi JN, Fei B, Roble S, Dota A, Katrapati P, Bezerra HG, et al. Assessment of coronary artery calcium using dual-energy subtraction digital radiography. Journal of digital imaging. 2012;25(1):129-36.
13. Song Y, Wu H, Wen D, Zhu B, Graner P, Ciancibello L, et al. Detection of coronary calcifications with dual energy chest X-rays: clinical evaluation. The International Journal of Cardiovascular Imaging. 2021;37(3):767-74.
14. Blaha MJ, Cainzos-Achirica M, Greenland P, McEvoy JW, Blankstein R, Budoff MJ, et al. Role of coronary artery calcium score of zero and other negative risk markers for cardiovascular disease: the Multi-Ethnic Study of Atherosclerosis (MESA). Circulation. 2016;133(9):849-58.
15. Peng AW, Dardari ZA, Blumenthal RS, Dzaye O, Obisesan OH, Iftekhar Uddin S, et al. Very high coronary artery calcium (≥ 1000) and association with cardiovascular disease events,

non–cardiovascular disease outcomes, and mortality: results from MESA. Circulation. 2021;143(16):1571-83.
16. Hubbell JH, Seltzer SM. Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z= 1 to 92 and 48 additional substances of dosimetric interest. National Inst. of Standards and Technology-PL, Gaithersburg, MD (United …, 1995.
17. Badal A, Badano A. Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit. Medical physics. 2009;36(11):4878-80.
18. Nasir K, Michos ED, Blumenthal RS, Raggi P. Detection of high-risk young adults and women by coronary calcium and National Cholesterol Education Program Panel III guidelines. Journal of the American College of Cardiology. 2005;46(10):1931-6.
19. Akosah KO, Schaper A, Cogbill C, Schoenfeld P. Preventing myocardial infarction in the young adult in the first place: how do the National Cholesterol Education Panel III guidelines perform? Journal of the American College of Cardiology. 2003;41(9):1475-9.
20. Müller K, Datta S, Ahmad M, Choi JH, Moore T, Pung L, et al. Interventional dual-energy imaging—Feasibility of rapid kV-switching on a C-arm CT system. Medical physics. 2016;43(10):5537-46.
21. Shi L, Lu M, Bennett NR, Shapiro E, Zhang J, Colbeth R, et al. Characterization and potential applications of a dual-layer flat-panel detector. Medical physics. 2020;47(8):3332-43.
22. Euler A, Nowak T, Bucher B, Eberhard M, Schmidt B, Flohr TG, et al. Assessment of bone mineral density from a computed tomography topogram of photon-counting detector computed tomography—effect of phantom size and tube voltage. Investigative Radiology. 2021;56(10):614-20.