Validation and Prognosis of Coronary Artery Calcium Scoring in Nontriggered Thoracic Computed Tomography: Systematic Review and Meta-analysis
Xueqian Xie, Yingru Zhao, Geertruida H. de Bock, Pim A. de Jong, Willem P. Mali, Matthijs Oudkerk and Rozemarijn Vliegenthart

Circ Cardiovasc Imaging. 2013;6:514-521; originally published online June 11, 2013; doi: 10.1161/CIRCIMAGING.113.000092

Circulation: Cardiovascular Imaging is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-9651. Online ISSN: 1942-0080

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circimaging.ahajournals.org/content/6/4/514

Data Supplement (unedited) at:
http://circimaging.ahajournals.org/content/suppl/2013/06/11/CIRCIMAGING.113.000092.DC1.html

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Imaging can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Imaging is online at:
http://circimaging.ahajournals.org//subscriptions/
Validation and Prognosis of Coronary Artery Calcium Scoring in Nontriggered Thoracic Computed Tomography

Systematic Review and Meta-analysis

Xueqian Xie, MD; Yingru Zhao, MD; Geertruida H. de Bock, PhD; Pim A. de Jong, MD, PhD; Willem P. Mali, MD, PhD; Matthijs Oudkerk, MD, PhD; Rozemarijn Vliegenthart, MD, PhD

Background—Coronary calcium score (CS), traditionally based on electrocardiography-triggered computed tomography (CT), predicts cardiovascular risk. Currently, nontriggered thoracic CT is extensively used, such as in lung cancer screening. The purpose of the study was to determine the correlation in CS between nontriggered and electrocardiography-triggered CT, and to evaluate the prognostic performance of the CS derived from nontriggered CT.

Methods and Results—PubMed, Embase, and Web of Knowledge were searched until November 2012. Two reviewers independently screened 2120 records to identify studies reporting the CS in nontriggered CT and extracted information. Study quality was evaluated by standardized assessment tools. Cohen κ was extracted for agreement of CS categories between nontriggered and electrocardiography-triggered CT (validation). Hazard ratio (HR) was extracted for prognostic performance. Five studies about validation comprising 1316 individuals were included. Five studies about prognosis comprising 34028 cardiac asymptomatic individuals, mainly from lung cancer screening trials, were included. All studies were of high quality. Meta-analysis could only be performed for validation studies because studies on prognostic performance were highly heterogeneous. Pooled Cohen κ for agreement between the 2 techniques was 0.89 (95% confidence interval, 0.83–0.95) for increasing CS categories. Increasing CS categories were associated with increasing risk of cardiovascular death or events. Nontriggered CT yielded false-negative CS in 8.8% of individuals and underestimated high CS in 19.1% of individuals.

Conclusions—Our analysis shows the prognostic value and potential role of nontriggered assessment of coronary calcium, but it does not suggest that electrocardiography-triggered CT should be replaced by nontriggered examinations. (Circ Cardiovasc Imaging. 2013;6:514-521.)

Key Words: calcium score □ computed tomography □ meta-analysis □ nontriggered cardiac imaging technique □ review, systematic

The online-only Data Supplement is available at http://circimaging.ahajournals.org/lookup/suppl/doi:10.1161/CIRCIMAGING.113.000092/-/DC1.

Correspondence to Rozemarijn Vliegenthart, MD, PhD, Center for Medical Imaging–NE Netherlands, Department of Radiology, University of Groningen, University Medical Center Groningen, Hanzeplein 1, 9700 RB Groningen, The Netherlands. E-mail r.vliegenthart@umcg.nl

© 2013 American Heart Association, Inc.
those at high cardiovascular risk, there may be an enormous unused primary prevention potential.\textsuperscript{11} Also, deriving the CS from the same examination as used in lung cancer screening may positively impact the cost-effectiveness of screening.

Because motion of coronary arteries influences calcium scoring,\textsuperscript{12} the use of coronary calcium scoring in nontriggered CT is still being debated.\textsuperscript{13} With the increasing interest in lung cancer screening, this is an optimal moment to investigate the potential use of nontriggered CT for calcium scoring. However, compared with the extensive publications in electrocardiography-triggered cardiac CT, the literature on calcium scoring in nontriggered thoracic CT is relatively limited. Therefore, we performed a systematic review and meta-analysis to investigate the validity and prognostic value of calcium scoring derived from nontriggered thoracic CT.

Methods

This study was performed according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA).\textsuperscript{14}

Information Sources and Search

We searched PubMed, Embase, and Web of Knowledge until November 2012, using terms related to CT, nontriggered thoracic examination, and coronary calcium without language restrictions (Table I in the online-only Data Supplement). Unpublished studies were not included.

Study Selection

Two reviewers (XQ.X. and YR.Z.) with ≥28 years of experience in thoracic and cardiovascular radiology participated in literature selection. Each record was evaluated independently. Disagreement in literature selection was resolved by consensus. Studies were included in the systematic review when they (1) evaluated cardiac asymptomatic adult humans or phantoms; (2) analyzed one of the following topics about calcium scoring in nontriggered CT: agreement between nontriggered and electrocardiography-triggered CT, or prognostic performance to predict death or events; (3) used ≥16-row multidetector CT as electrocardiography-triggered examination when electrocardiography-triggered CT was used as reference examination. Sixteen-row multidetector CT was used as minimum CT generation because previous research showed higher accuracy and reproducibility in calcium scoring for 16-row multidetector CT compared with earlier generation CT systems.\textsuperscript{15}

Articles were excluded when they (1) were reviews, abstracts, case reports, or letters; (2) investigated participants with confounding factors, for example, pacemaker or defibrillator implant, and cardiac surgery. When multiple similar publications based on the same trial were identified, only the study with the largest sample size was included to avoid possible duplicate reporting.

Subsequently, meta-analysis was performed in studies on agreement between nontriggered and electrocardiography-triggered CT, when the studies used the same calcium scoring method, that is, continuous CS and 4 CS categories (0, 1–99, 100–399, and ≥400). No meta-analysis could be performed on the studies on prognostic value because of large heterogeneity in calcium quantification methods, CS categorization, and outcomes.

Data Collection Process

A standardized data extraction form was used to collect study and participant characteristics, methodology, and main results. Two reviewers (XQ.X. and YR.Z.) collected data independently. Disagreement in data collection was resolved by consensus.

For results of studies on agreement of calcium scoring between nontriggered and electrocardiography-triggered CT, a correlation coefficient was extracted for continuous data, and Cohen κ and concordance percentage were extracted for categorical data. When available, the subject number with CS of >0, <400, and ≥400 was extracted for the 2 techniques. A CS of ≥400 is commonly considered as indicating high cardiovascular risk.\textsuperscript{3, 5} Therefore, the diagnostic performance of nontriggered CT was calculated using electrocardiography-triggered CT as reference. The percentage of false-negative CS was calculated as the percentage of subjects with zero CS in nontriggered CT among subjects with CS≥0 in electrocardiography-triggered CT. The percentage of underestimated high-risk CS was considered as the percentage of subjects with CS<400 in nontriggered CT among subjects with CS≥400 in electrocardiography-triggered CT. The percentage of overestimated high-risk CS was calculated as the percentage of subjects with CS≥400 in nontriggered CT among subjects with CS<400 in electrocardiography-triggered CT.

For prognostic performance of calcium scoring in nontriggered CT, HR for increasing CS categories derived from nontriggered CT to predict cardiovascular death or cardiovascular events (coronary heart disease, cerebrovascular disease, heart failure, peripheral arterial disease, aortic aneurysm, etc) was extracted. When possible, unadjusted and adjusted HR with 95% confidence interval was extracted. Furthermore, the number of subjects with zero CS was extracted, as well as the number of subsequent cardiovascular deaths or events among these subjects.

Study Quality Assessment

Two reviewers (XQ.X. and YR.Z.) evaluated study quality independently on the studies included in the systematic review. Disagreement in quality assessment was resolved by consensus. Two quality assessment tools for different type of study were used to evaluate methodological quality and potential sources of bias, as described next.

For validation studies on agreement between nontriggered and electrocardiography-triggered CT, the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) with 14 standard items was used.\textsuperscript{16} For each study, a quality score was derived by assigning 1 point to each fulfilled item, 0.5 to an unclear item, and 0 to an unmet item, with a total possible score of 14 (Table II in the online-only Data Supplement).

For prognostic studies, the quality assessment criteria to evaluate reports on prognosis of coronary artery calcium (CAC) in the American College of Cardiology/Foundation/American Heart Association (ACC/F/AHA) guideline with 8 standard criteria were used.\textsuperscript{17} For each study, a quality score was derived by assigning 0 to 3 points to each criterion, with a total possible score of 16 (Table III in the online-only Data Supplement).

Synthesis of Results and Risk of Bias

Pooling calculations of agreement between the 2 techniques were performed using the Hedges–Veeva random effects model and Z test for overall effect. Pooling calculation was performed if there were ≥2 studies reporting the same measurement. Heterogeneity was tested using Q statistic and I² index. A 2-sided P value for Q statistic <0.10 or I²>50% was considered to indicate heterogeneity. The random effects model was used regardless of the heterogeneity test, although results in Q statistics and I² index were still stated. Publication bias was evaluated with the Begg and Mazumdar rank correlation and Egger regression test if the number of effect size in the included studies was ≥3. For other statistical analysis, a 2-sided P value <0.05 was considered as significant. Statistical analysis was performed using R version 2.14.2 (R Foundation, Vienna, Austria) and Stata version 11.0 (Statacorp LP, College Station, TX).

Results

Study Selection

The search of the 3 databases elicited 2120 records after removal of duplicate records. Ten studies were included in systematic review, including 5 on agreement between nontriggered and electrocardiography-triggered CT,\textsuperscript{17–21} and 5 on prognostic performance.\textsuperscript{10, 22–25} Subsequently, meta-analysis was performed in 3 studies\textsuperscript{20, 22, 24} with consistent methodology.
on agreement between nontriggered and electrocardiography-triggered CT (Figure 1).

Study Characteristics

The systematic review included 35,344 participants (range of mean age, 51–65 years), comprising 21,558 (61%) men, 13,736 (39%) women, and 50 (0.1%) individuals without indicated sex (Tables 1 and 2). Six (60%) studies were prospective, and 4 (40%) studies were retrospective. Four studies (40%) were from North America, 3 (30%) from Europe, and 3 studies (30%) from Asia. All studies were published in English.

Different CT modalities were used, ranging from single-slice to 64-row multidetector CT. Also CT systems as part of single-photon emission CT/CT and positron emission tomography/CT were used. Low-dose acquisition was applied in 8 studies (80%), and normal dose was applied in 1 study (10%). In 1 (10%) study, the radiation dose was not reported. Scan data derived from nontriggered CT were reconstructed with different slice thicknesses, ranging from 1.25 to 10 mm (Tables 1 and 2). The most commonly used slice thicknesses were 2.5/3 mm and 5 mm. Four studies used a (medium-) smooth kernel, the other studies did not indicate the applied kernel. Six studies used Agatston scoring, whereas 4 others used visual grading of the presence and extent of coronary calcification. No study used contrast media. No phantom study was included.

Study Quality

All 5 studies on agreement between the nontriggered and electrocardiography-triggered CT were of high quality (score ≥10 according to the QUADAS tool). Suboptimal scores were present in 2 QUADAS items: no study mentioned uninterpretable results (item 13); 3 studies did not mention whether there were withdrawals (item 14; Table II in the online-only Data Supplement).

All 5 studies on prognostic performance were of high quality (a score ≥12 according the quality assessment criteria to evaluate prognosis of coronary calcification). Suboptimal scores were present for criterion 4: no study reported results by ethnicity (Table III in the online-only Data Supplement).

Validation of Calcium Scoring in Nontriggered CT

Five studies were included in the systematic review, comprising 1316 cardiac asymptomatic participants (Table 1). Diagnostic performance of nontriggered CT was calculated in 4 studies with 1153 subjects (Table 3), in which, 137 (11.9%) subjects had CS of 100 to 400 in nontriggered CT. Fifty-five subjects (8.8%) showed no coronary calcification in nontriggered CT examination among 625 subjects with CS>0 in electrocardiography-triggered CT. In those 55 subjects, 52 (8.3%) had CS 1 to 100 in electrocardiography-triggered CT, and 3 (0.5%) subjects had CS 100 to 400. Among 162 subjects

---

Table 1. Characteristics of Studies on Agreement of Calcium Scoring Between Nontriggered and Triggered CT

<table>
<thead>
<tr>
<th>Studies, y</th>
<th>Patients, n</th>
<th>Men, %</th>
<th>Age, y±SD</th>
<th>Setting of Study</th>
<th>Type of Participants</th>
<th>CT Type</th>
<th>Radiation Dose Setting</th>
<th>Nontriggered CT</th>
<th>Electrocardiography-Triggered CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budoff 2011</td>
<td>50</td>
<td>n/a</td>
<td>n/a</td>
<td>COPD cohort</td>
<td>Invited smokers of &gt;10 pack-years</td>
<td>64-MDCT</td>
<td>Low</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Einstein 2010</td>
<td>492</td>
<td>44</td>
<td>n/a</td>
<td>Routine clinical population</td>
<td>Consecutively referred adults</td>
<td>16-SPECT/CT</td>
<td>Low</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Kim 2008</td>
<td>128</td>
<td>100</td>
<td>52±7</td>
<td>Lung cancer screening</td>
<td>Consecutively referred elder smokers</td>
<td>40-MDCT</td>
<td>Low</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Kirsch 2011</td>
<td>163</td>
<td>78</td>
<td>51±9</td>
<td>Asymptomatic</td>
<td>Consecutively referred elder adults</td>
<td>16- and 64-MDCT</td>
<td>Normal</td>
<td>5.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Wu 2008</td>
<td>483</td>
<td>66</td>
<td>62±13</td>
<td>Lung cancer screening</td>
<td>Self-referred elder adults</td>
<td>16-MDCT</td>
<td>Low</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

COPD indicates chronic obstructive pulmonary disease; CT, computed tomography; MDCT, multidetector computed tomography; n/a, not available; PET, positron emission tomography; and SPECT, single-photon emission computed tomography.
with CS≥400 in electrocardiography-triggered CT, nontriggered CT underestimated the CS in 31 subjects (19.1%). In these 31 subjects, 2 (1.2%) had CS 1 to 100 in nontriggered CT, and 29 (17.9%) subjects had CS 100 to 400. However, among 991 subjects with CS<400 in electrocardiography-triggered CT, nontriggered CT underestimated the CS in 31 subjects (19.1%). In those 31 subjects, 2 (1.2%) had CS 1 to 100 in nontriggered CT and, thus, overestimated the CS. In those 26 subjects, 1 (0.1%) had CS 1 to 100 in electrocardiography-triggered CT, and 25 (2.5%) subjects had CS 100 to 400.

Meta-analysis was performed in 3 studies comprising 661 participants (Figure 2).17,19,21 The study by Kirsch et al20 could not be included because it evaluated the amount of coronary calcification using visual grading score. The pooled correlation coefficient for CS was 0.94 (95% confidence interval, 0.89–0.97). The pooled Cohen κ was 0.89 (95% confidence interval, 0.83–0.95) for 4 categories of the CS. Heterogeneity was found in the pooling calculation of the CS (P for Q statistic <0.001 and I² >50%). No publication bias was found in the pooling calculation of the CS (P>0.05). Publication bias testing was not performed in the pooling calculation of 4 CS categories because of insufficient number of studies.

### Prognosis of Calcium Scoring in Nontriggered CT

Five studies were included, comprising 34028 cardiac asymptomatic participants (Table 2).10,22–25 In the 5 studies, mean follow-up duration was 45 months (range, 10–72 months). None of the participants in the included studies had a history or symptoms of cardiovascular diseases before CT examination.

### Table 2. Characteristics of Studies on Prognostic Performance of Calcium Scoring for Cardiovascular Death or Events

<table>
<thead>
<tr>
<th>Studies, y</th>
<th>Patients, n</th>
<th>Men, %</th>
<th>Age, y±SD</th>
<th>Setting of Study</th>
<th>Type of Participants</th>
<th>CT Type</th>
<th>Radiation Dose Setting</th>
<th>Slice Thickness, mm</th>
<th>Cohort Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itani 2004</td>
<td>6120</td>
<td>55</td>
<td>61±11</td>
<td>Lung cancer screening</td>
<td>Invited elder adults</td>
<td>Single-slice CT</td>
<td>Low</td>
<td>10</td>
<td>Nagano</td>
</tr>
<tr>
<td>Jacobs 2011</td>
<td>10410</td>
<td>58</td>
<td>62±12</td>
<td>Routine clinical population</td>
<td>Retrospectively included elder adults</td>
<td>4-, 8-, 16-, 40- and 64-MDCT</td>
<td>n/a</td>
<td>3.0–10</td>
<td>PROVIDI</td>
</tr>
<tr>
<td>Jacobs 2012</td>
<td>7557</td>
<td>83</td>
<td>60±6</td>
<td>Lung cancer screening</td>
<td>Invited elder healthy smokers</td>
<td>16-MDCT</td>
<td>Low</td>
<td>3.1</td>
<td>NELSON</td>
</tr>
<tr>
<td>Shemesh 2010</td>
<td>8782</td>
<td>49</td>
<td>65±7</td>
<td>Lung cancer screening</td>
<td>Invited elder heavy smokers</td>
<td>Single- and multislice CT</td>
<td>Low</td>
<td>1.25–5</td>
<td>I-ELCAP</td>
</tr>
<tr>
<td>Sverzellati 2012</td>
<td>1159</td>
<td>68</td>
<td>58±6</td>
<td>Lung cancer screening</td>
<td>Invited elder heavy smokers</td>
<td>16-MDCT</td>
<td>Low</td>
<td>5</td>
<td>MILD</td>
</tr>
</tbody>
</table>

CT indicates computed tomography; I-ELCAP, international early lung cancer action program; MDCT, multidetector computed tomography; MILD, multicentric Italian lung detection; n/a, not available; NELSON, the Dutch–Belgian randomized lung cancer screening trial; and PROVIDI, the prognostic value of unrequested information in diagnostic imaging.

### Table 3. Agreement of Coronary Calcium Scoring Between Nontriggered Thoracic and Electrocardiography-Triggered Cardiac CT, and Diagnostic Performance of Calcium Scoring in Nontriggered CT

<table>
<thead>
<tr>
<th>Studies, y</th>
<th>Scoring in Nontriggered CT</th>
<th>Reference Scoring in Triggered CT</th>
<th>Agreement Between Nontriggered and Triggered CT</th>
<th>Diagnostic Performance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budoff 2011</td>
<td>CS</td>
<td>CS</td>
<td>n=0.96</td>
<td>False-Negative Calcium Score, %</td>
</tr>
<tr>
<td></td>
<td>4 categories of CS†</td>
<td>4 categories of CS†</td>
<td>κ=0.90, concordance=94%</td>
<td>Underestimated High Calcium Score, %</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Overestimated High Calcium Score, %</td>
</tr>
<tr>
<td>Einstein 2010</td>
<td>6 categories of CS‡</td>
<td>6 categories of CS‡</td>
<td>κ=0.89, concordance=63%</td>
<td>14.0</td>
</tr>
<tr>
<td>Kim 2008¹⁹</td>
<td>CS</td>
<td>CS</td>
<td>n=0.89</td>
<td>9.3</td>
</tr>
<tr>
<td>Kirsch 2011</td>
<td>Visual grading score§</td>
<td>CS</td>
<td>n=0.83</td>
<td>n/a</td>
</tr>
<tr>
<td>Wu 2008²⁰</td>
<td>4 categories of CS†</td>
<td>4 categories of CS†</td>
<td>κ=0.89, concordance=93%</td>
<td>2.3</td>
</tr>
</tbody>
</table>

CS indicates calcium score; CT, computed tomography; and n/c, not calculated because different scoring systems were used in nontriggered and electrocardiography-triggered CT.

*False-negative calcium score is indicated as percentage of CS=0 in nontriggered CT among those with CS>0 in triggered CT. Underestimated high-risk calcium score is indicated as percentage of CS<400 in nontriggered CT among those with CS≥400 in triggered CT. Overestimated high-risk calcium score is indicated as percentage of CS≥400 in nontriggered CT among those with CS<400 in triggered CT.

†Four categories of CS defined as 0–1, 1, 100–399, and ≥400.
‡Six categories of CS, defined as 0, 1–99, 100–399, 400–999, and ≥1000.
§Score assigned to each major coronary artery. 0, no calcification; 1, single pixel calcification; 3, dense calcification with blooming artifact; 2, calcification between 1 and 3. The visual grading score (range, 0–12) was calculated by summing the score for each artery.
During follow-up, 207 cardiovascular deaths and 675 cardiovascular events were observed. Overall, with increasing CS categories, increasing unadjusted and adjusted HR for cardiovascular death or events was observed. Risks in CS categories were not consistently reported; however, in 1 study, unadjusted and adjusted HR increased up to 7.5 and 5.3 for CS ≥1000, respectively (Table 4). A small percentage of subjects with zero CS in nontriggered CT had cardiovascular death or events. During a mean follow-up of 45 months, 47 cardiovascular deaths (0.55%) were found in 8487 subjects with zero CS, whereas 72 cardiovascular deaths (2.5%) were found in 6415 subjects with positive CS. However, the event rate for subjects with positive CS was higher. During follow-up, 160 cardiovascular deaths (2.5%) were found in 6415 subjects with positive CS, whereas 570 cardiovascular events (4.5%) occurred in 12,718 subjects with positive CS.

**Discussion**

In this systematic review and meta-analysis, we aimed to investigate whether coronary calcium scoring can be performed in nontriggered thoracic CT, for instance, used in lung cancer screening. A strong correlation in CS categories between nontriggered and electrocardiography-triggered CT was found. In cardiac asymptomatic elderly and smokers, mainly from lung cancer screening trials, increasing coronary calcium burden translated into a higher risk of cardiovascular death or events.

CS for individual atherosclerotic lesions is greatly influenced by motion. Regardless, we found that the correlation in CS between nontriggered and electrocardiography-triggered CT was excellent (r=0.94) on a group level. In broad CS categories, we found a high agreement between nontriggered and electrocardiography-triggered CT. However, event rate for subjects with positive CS was higher. During follow-up, 160 cardiovascular deaths (2.5%) were found in 6415 subjects with positive CS, whereas 570 cardiovascular events (4.5%) occurred in 12,718 subjects with positive CS.

We found that nontriggered CT can yield a false-negative CS in 9% of individuals compared with electrocardiography-triggered CT. Furthermore, we found that a zero CS in nontriggered CT indicates a low cardiovascular risk, although nontriggered CT cannot reliably exclude coronary calcification. When a high CS (≥400) is found in asymptomatic individuals, the risk of cardiovascular events is elevated. The ACCF/AHA consensus document suggests to consider these individuals as candidates for intensive preventive therapies. The probability of overestimating the CS is low and, thus, it is reasonable to assume an elevated cardiovascular risk in case of a CS ≥400 in nontriggered CT. However, nontriggered CT underestimated the CS in a non-negligible percentage of individuals with CS ≥400 in electrocardiography-triggered CT, thus, underestimating cardiovascular risk. In the validation study, 11.9% had a CS of 100 to 400 in the nontriggered CT examination. In this relatively small proportion of the included study populations, dedicated electrocardiography-triggered CT could be considered, to assess whether the CS is actually ≥400. This proportion is much lower than the population percentage in which calcium scoring could be considered according to current consensus documents (40%).

In this study, HRs of CS categories for cardiovascular events were generally lower than in a previous systematic review on calcium scoring derived from electrocardiography-triggered CT. For example, adjusted HR for cardiovascular events was up to 5.3 for CS ≥1000 in our study, <10.8 in a previous report in electrocardiography-triggered CT in an elderly population. The relative risk is usually based on the risk of subjects without coronary calcium at baseline as reference category. During a mean follow-up of 45 months, we found that 1.3% subjects without coronary calcium had a cardiovascular event. In contrast, a meta-analysis by Sarwar et al on electrocardiography-triggered CT reported only 0.47% of subjects without coronary calcium had a cardiovascular event during a mean follow-up of 50 months. In that meta-analysis, studies mainly consisted of middle-aged individuals at low-to-intermediate cardiovascular risk, referred for cardiovascular risk evaluation. In contrast, the majority of the populations in the prognostic studies on calcium scoring using nontriggered CT comprised participants of lung cancer screening trials. The generally higher age and heavier smoking history in the prognostic studies included in our study likely at least partly explain the higher event rate in individuals with zero CS.
Besides, this higher event rate in case of zero CS for nontriggered CT may also be explained by the fact that a proportion of the individuals without coronary calcification on the nontriggered CT, actually have a positive CS in electrocardiography-triggered CT. For example, in 483 subjects, Wu et al.21 reported an interobserver variability of calcium scoring in nontriggered CT, with a mean correlation coefficient, 0.77–0.91; intraclass correlation coefficient, 0.93–0.99.10,18,20,21,23–25

The agreement of repeated calcium scoring in nontriggered CT within and between observers is high, although slightly lower than in electrocardiography-triggered CT. Nearly all studies in this systematic review investigated either intraobserver or interobserver variability of calcium scoring in nontriggered CT. For example, in 483 subjects, Wu et al reported an interobserver variability of 3.6% for electrocardiography-triggered CT, and of 9.6% for nontriggered CT. However, all studies found a very strong concordance in score categorization within and between observers (κ = 0.77–0.91; intraclass correlation coefficient, 0.93–0.99).10,18,20,21,23–25

The majority of included studies (80%) were based on low-dose thoracic CT,10,17–19,21,22,24,25 which has a lower radiation dose than a dedicated cardiac CT for calcium scoring. A typical effective radiation dose for low-dose CT used in lung cancer screening is 0.8 to 0.9 mSv for normal sized body.8,21 However, the mean dose for a cardiac CT for calcium scoring is ≈ 1.0 to 2.9 mSv, depending on scanner type and scanning protocol.23,31

**Clinical Implication**

A large number of nontriggered CT examinations are annually performed worldwide. In the aging and smoking population, coronary calcification is a common finding. A lung cancer screening trial reported that >70% of the participants had coronary calcification.10 The group at risk for lung cancer

---

**Table 4. Prognostic Performance of Coronary Calcium Scoring for Cardiovascular Death or Events in Nontriggered CT**

<table>
<thead>
<tr>
<th>Studies, y</th>
<th>Follow-Up, mo (range)</th>
<th>End Point, n</th>
<th>Calcium Scoring Method in Nontriggered CT</th>
<th>Calcium Scoring Cutoff</th>
<th>Event Number/Category</th>
<th>Unadjusted Hazard Ratio of Calcium Score (95% CI)</th>
<th>Adjusted Hazard Ratio of Calcium Score (95% CI)</th>
<th>Adjusted for Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itani 200422</td>
<td>48</td>
<td>Cardiac death, 14</td>
<td>Presence of coronary calcification</td>
<td>Zero calcium score Positive calcium score</td>
<td>4/4914, 0.08</td>
<td>1.0 (reference)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Jacobs 201123</td>
<td>18</td>
<td>Cardiovascular event, 515</td>
<td>4 categories of visual grading score*</td>
<td>Visual score: 0 Visual score: 1–2 Visual score: 3–5 Visual score: 6–12</td>
<td>62/3435, 1.8 113/2498, 4.5 149/2603, 5.7 191/1874, 10.2</td>
<td>1.0 (reference) 2.8 (2.0–3.8) 3.8 (2.9–5.2) 6.9 (5.2–9.2)</td>
<td>1.0 (reference) 2.2 (1.6–3.0) 2.5 (1.8–3.4) 3.7 (2.7–5.2)</td>
<td>Age, sex, smoking status, hypertension, hypercholesterolemia, and diabetes mellitus</td>
</tr>
<tr>
<td>Jacobs 201224</td>
<td>10 (1–21)</td>
<td>Cardiovascular event, 127</td>
<td>4 categories of calcium score</td>
<td>Calcium score: 0 Calcium score: 1–100 Calcium score: 101–1000 Calcium score: &gt;1000</td>
<td>10/1814, 0.6 27/2191, 1.2 32/2267, 1.4 58/1285, 4.5</td>
<td>1.0 (reference) 1.9 (0.9–4.2) 2.2 (1.0–4.8) 7.5 (3.6–15.7)</td>
<td>1.0 (reference) 1.8 (0.8–3.9) 1.9 (0.9–4.2) 5.3 (2.5–11.6)</td>
<td>Age, sex, smoking status, hypertension, hypercholesterolemia, and diabetes mellitus</td>
</tr>
<tr>
<td>Shemes 201025</td>
<td>72 (1–92)</td>
<td>Cardiovascular death, 193</td>
<td>3 categories of visual grading score†</td>
<td>Visual score: 0 Visual score: 1–3 Visual score: 4–12</td>
<td>43/3573, 1.2 66/3569, 1.8 84/1640, 5.1</td>
<td>1.0 (reference) 1.6 (1.1–2.4) 4.7 (3.6–6.8)</td>
<td>1.0 (reference) 1.0 (0.7–1.5) 2.1 (1.4–3.1)</td>
<td>Age, sex, smoking pack-years</td>
</tr>
<tr>
<td>Sverzellati 201226</td>
<td>36 (1–54)</td>
<td>Cardiovascular event, 33</td>
<td>2 categories of calcium score</td>
<td>Calcium score: ≤400 Calcium score: &gt;400</td>
<td>26/1079, 2.4 7/80, 8.8</td>
<td>1.0 (reference) n/a</td>
<td>2.9 (1.1–7.3)</td>
<td>Age, sex, diabetes mellitus, hypertension, smoking status, and smoking duration</td>
</tr>
</tbody>
</table>

CI indicates confidence interval; CT, computed tomography; and n/a, not available.

*Grading score assigned to each major coronary artery. 0, no calcification; 1, 1–2 calcifications; 2, >2 calcifications, or 1 calcification extending >2 slices; 3, greater than or equal to three artery length showing calcification; 2, one third to two thirds; 3, greater than or equal to two thirds. Three visual grades were stratified by the sum of the score (0, 1–3, and 4–12).
overlaps with the group at highest risk of cardiovascular diseases because at least aging and smoking are 2 major risk factors for both diseases. There may be an enormous primary prevention potential if the CS can be derived from the same examination, at least in participants of lung cancer screening trials. Although results from the 1 study in a clinical population suggest that the extent of coronary calcification is also predictive outside lung cancer screening setting, more studies are needed to confirm the value of calcium scoring in routine clinical thoracic CT.

We observed that CS categorization between nontriggered and electrocardiography-triggered CT correlated very well, and increasing CS categories based on nontriggered CT are predictive of increasing cardiovascular risk. Thus, for subjects who were examined by nontriggered thoracic CT, the cardiovascular risk could potentially be stratified by performing calcium scoring. Subjects identified in nontriggered CT as having high CS could be considered as candidates of intensive risk factor modification, especially in an aging and smoking population, such as the participants in lung cancer screening. However, a zero CS in nontriggered CT does not exclude coronary calcification.

Furthermore, cardiovascular event rate of subjects without CS in nontriggered CT is higher than in electrocardiography-triggered CT. Absent coronary calcification in nontriggered CT may not reliably exclude the risk of cardiovascular events. Future studies on this topic are needed to provide stronger support for coronary calcium scoring in nontriggered CT.

**Limitations**

First, despite our favorable results it remains to be clarified whether differences in the accuracy between nontriggered and electrocardiography-triggered CS measures translate into differences in prognostic value. For example, a zero CS in nontriggered CT may render a positive CS in electrocardiography-triggered examinations. Second, for agreement in calcium scoring between nontriggered and electrocardiography-triggered CT, the number of studies and participants in the meta-analysis was relatively low. To compare calcium scoring between nontriggered and electrocardiography-triggered CT, the patients have to be scanned twice in a short time interval, and at doubled radiation dose. This could contribute to the relatively small number of studies in this field. Our conclusions are not only based on the pooling calculations, but also on the systematic review. Besides, for the second part of our study, about the prognostic value, the aggregated sample size was >30000 individuals. Third, different calcium scoring methods were used in studies on prognostic performance. Therefore, meta-analysis could not be performed to assess predictive value of the CS derived from nontriggered CT. However, heavier calcium burden was in the systematic review associated with increasing cardiovascular risk. Finally, included studies were fairly heterogeneous in terms of participant population, imaging equipment, and acquisition protocol. Those factors weakened the strength of meta-analysis. We used a random effects model to compensate for at least some of the heterogeneity in the pooling calculation. However, the differences in imaging procedures also reflect the heterogeneity of procedures in clinical practice. Despite different CT equipment and calcium scoring methods, at least the presence or absence of coronary calcium is clear. Results on the presence and absence of coronary calcification should not differ significantly based on important CS categories. Thus, our conclusions based on findings related to those points are solid.

**Conclusions**

In this systematic review and meta-analysis, strong agreement in CS categorization was found between nontriggered CT and electrocardiography-triggered CT. Compared with electrocardiography-triggered CT, a high CS category in nontriggered CT is a fairly reliable finding. However, nontriggered CT yielded false-negative CS in 8.8% of individuals, and underestimated high CS in 19.1%. In cardiac asymptomatic participants mainly from lung cancer screening trials, increasing CS categories in nontriggered CT were associated with increasing risk of cardiovascular events. Our analysis presents preliminary evidence for the prognostic value and potential role of calcium scoring in nontriggered CT. However, it does not suggest that nontriggered examinations can replace dedicated, electrocardiography-triggered CT.

**Acknowledgments**

The authors wish to thank Dr Estelle J.K. Noach for revision of the article.

**Disclosures**

None.

**References**

In this study, coronary calcium score categories based on nontriggered computed tomography (CT) were fairly comparable with electrocardiography-triggered CT results. Thus, cardiac asymptomatic individuals, such as those taking part in lung cancer screening trials, can be stratified into broad categories of cardiovascular risk based on one and the same CT examination. Calcium scoring derived from nontriggered thoracic CT has great potential to identify individuals with heavy coronary calcium burden, who are at increased cardiovascular risk and who could qualify for stricter cardiovascular risk factor control. Despite our favorable results, it remains to be clarified whether differences in the accuracy between nontriggered and electrocardiography-triggered calcium score measures translate into differences in prognostic value. For example, a zero calcium score in nontriggered CT may render a positive calcium score in electrocardiography-triggered examinations.
SUPPLEMENTAL MATERIAL

Online-only Data Supplement Table A. Literature search strategy

PubMed

("Tomography, X-Ray Computed"[MeSH] OR "computed tomography"[tiab] OR CT[tiab] OR “MDCT”) AND ("untriggered" OR “ungated” OR “non-gated” OR “non-triggered” OR “non-electrocardiogram” OR “thorax”[MeSH] OR “chest” OR “thoracic” OR “lung” OR “pulmonary” OR “torso”) AND (“coronary vessels”[MeSH] OR "Coronary") AND (“Calcium” OR “calcification” OR “calcific” OR “calcified”) AND 1900/01:2012/11[dp]

EmBase

#1: ((Computed tomography) OR CT:ab,ti OR MDCT) AND (untriggered OR ungated OR non-gated OR non-triggered OR non-electrocardiogram OR thorax OR chest OR thoracic OR lung:ab,ti OR pulmonary:ab,ti OR torso) AND Coronary AND (Calcium OR calcification OR calcific OR calcified)

Grammar in advanced search: #1 AND [1-1-1900]/sd NOT (#1 AND [30-11-2012]/sd)

Web of Knowledge

#1 topic: ((Computed tomography) OR CT OR MDCT)

#2 topic: (untriggered OR ungated OR non-gated OR non-triggered OR non-electrocardiogram OR thorax OR chest OR thoracic OR lung OR pulmonary OR torso)

#3 topic: Coronary

#4 topic: (Calcium OR calcification OR calcific OR calcified)

Grammar: #1 topic and #2 topic and #3 topic and #4 topic
Online-only Data Supplement Table B. Quality assessment for validation studies on agreement and diagnostic performance, by the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool

<table>
<thead>
<tr>
<th>Item</th>
<th>Budoff 2011¹</th>
<th>Einstein 2010²</th>
<th>Kim 2008³</th>
<th>Kirsch 2011⁴</th>
<th>Wu 2008⁵</th>
<th>Score†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Representative patient sample</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>12.0</td>
</tr>
<tr>
<td>2. Selection criteria clearly described</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>12.0</td>
</tr>
<tr>
<td>3. Acceptable reference standard</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>12.5</td>
</tr>
<tr>
<td>4. Acceptable delay between tests*</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.0</td>
</tr>
<tr>
<td>5. Partial verification avoided</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>6. Differential verification avoided</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>7. Incorporation avoided</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>8. Adequate index test description</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>10. Index test blinded to reference standard</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>11. Reference standard blinded to index test</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>12. Clinical data available as in practice</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>13. Uninterpretable test results reported</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
<tr>
<td>14. Withdrawals explained</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>13.5</td>
</tr>
</tbody>
</table>

*+“ = yes; “-” = no; empty = unclear

*Maximum delay of 2 months between non-triggered and reference examination was considered as acceptable

†For each study, a quality score was accumulated by assigning 1 point to “yes” item, 0.5 point to “unclear” item, and 0 to “no” item. The total possible score was 14. A score of ≥10 points was considered as high quality, and a score between 6 and 9 points as moderate quality, a score of ≤5 as low quality.
Online-only Data Supplement Table C. Quality assessment for studies on prognosis, by the quality assessment criteria of prognostic studies on coronary artery calcium in American College of Cardiology Foundation / American Heart Association (ACCF/AHA) guideline

<table>
<thead>
<tr>
<th>Study</th>
<th>Criterion 1: Retrospective vs. prospective study</th>
<th>Criterion 2: Potential for referral bias</th>
<th>Criterion 3: Reporting coronary calcification by CHD death or myocardial infarction</th>
<th>Criterion 4: Reporting of results by gender or ethnicity</th>
<th>Criterion 5: Sample size greater than 1000</th>
<th>Criterion 6: Potential for limited challenge</th>
<th>Criterion 7: Risk factor reporting</th>
<th>Criterion 8: Covariate or risk-adjusted outcomes</th>
<th>Score*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itani 2004(^6)</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Jacobs 2011(^7)</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Jacobs 2012(^8)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Shemesh 2010(^9)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Sverzellati 2012(^10)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>14</td>
</tr>
</tbody>
</table>

CHD = coronary heart disease.

*For each study, a quality score was accumulated by assigning a score for each criterion as the following:

- Criterion 1: Retrospective vs. prospective study (1=retrospective, 2=prospective).
- Criterion 2: Potential for referral bias (0=clinically referred patients, 1=unselected cohort, 2=population sample).
- Criterion 3: Reporting coronary calcification by CHD death or myocardial infarction (1=no, 2=yes).
• Criterion 4: Reporting of results by gender or ethnicity (0=no, 1=gender only, 2=ethnicity only, 3=both).

• Criterion 5: Sample size ≥ 1000 (0=no, 1=yes).

• Criterion 6: Potential for limited challenge (1=no reporting of calcium outcomes in low- to high-risk global risk scores, 2=reporting of calcium outcomes in low- to high-risk global risk scores).

• Criterion 7: Risk factor reporting (1=historical only, 2=measured in subset, 3=measured in all subjects).

• Criterion 8: Covariate or risk-adjusted outcomes (0=no, 1=yes).

Total possible score was 16. A score of ≥11 points was considered as high quality, and a score between 7 and 10 points as moderate quality, a score of ≤6 as low quality.
References


