



Published in final edited form as:

*Radiology*. 2015 July ; 276(1): 82–90. doi:10.1148/radiol.15142062.

## Association of Coronary Artery Calcification and Mortality in the National Lung Screening Trial:

### A Comparison of Three Scoring Methods

**Caroline Chiles, MD, Fenghai Duan, PhD, Gregory W. Gladish, MD, James G. Ravenel, MD, Scott G. Baginski, MD, Bradley S. Snyder, MS, Sarah DeMello, MS, Stephanie S. Desjardins, BS, Reginald F. Munden, MD, DMD, MBA, and For the NLST Study Team**

Department of Radiology, Wake Forest University Health Sciences Center, Medical Center Boulevard, Winston-Salem, NC 27157 (C.C.); Department of Biostatistics and Center for Statistical Sciences, Brown University School of Public Health, Providence, RI (F.D., B.S.S., S.D.); Department of Diagnostic Radiology, University of Texas MD Anderson Cancer Center, Houston, Tex (G.W.G.); Department of Radiology, Medical University of South Carolina, Charleston, SC (J.G.R.); Summit Radiology LLC, Oconomowoc, Wis (S.G.B.); American College of Radiology Imaging Network, Philadelphia, Pa (S.S.D.); and Department of Radiology, Houston Methodist Hospital and Research Institute, Houston, Tex (R.F.M.)

### Abstract

**Purpose**—To evaluate three coronary artery calcification (CAC) scoring methods to assess risk of coronary heart disease (CHD) death and all-cause mortality in National Lung Screening Trial (NLST) participants across levels of CAC scores.

**Materials and Methods**—The NLST was approved by the institutional review board at each participating institution, and informed consent was obtained from all participants. Image review was HIPAA compliant. Five cardiothoracic radiologists evaluated 1575 low-dose computed tomographic (CT) scans from three groups: 210 CHD deaths, 315 deaths not from CHD, and 1050 participants who were alive at conclusion of the trial. Radiologists used three scoring methods: overall visual assessment, segmented vessel-specific scoring, and Agatston scoring. Weighted Cox proportional hazards models were fit to evaluate the association between scoring methods and outcomes.

---

**Address correspondence to** C.C. (cchiles@wakehealth.edu).

#### **Author contributions:**

Guarantors of integrity of entire study, C.C., R.F.M.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, C.C., F.D., G.W.G., J.G.R., S.G.B., S.D., R.F.M.; clinical studies, C.C., F.D., G.W.G., J.G.R., S.S.D., R.F.M.; experimental studies, C.C., F.D., S.G.B., R.F.M.; statistical analysis, F.D., B.S.S., S.D., S.S.D.; and manuscript editing, C.C., F.D., G.W.G., J.G.R., S.G.B., B.S.S., S.D., R.F.M.

**Disclosures of Conflicts of Interest:** C.C. Financial activities related to the present article: disclosed no relevant relationships. Financial activities not related to the present article: author received money for consultancies from the National Cancer Institute and the American College of Radiology Imaging Network. Other relationships: disclosed no relevant relationships. **F.D.** disclosed no relevant relationships. **G.W.G.** disclosed no relevant relationships. **J.G.R.** disclosed no relevant relationships. **S.G.B.** disclosed no relevant relationships. **B.S.S.** disclosed no relevant relationships. **S.D.** disclosed no relevant relationships. **S.S.D.** disclosed no relevant relationships. **R.F.M.** disclosed no relevant relationships.

**Results**—In multivariate analysis of time to CHD death, Agatston scores of 1–100, 101–1000, and greater than 1000 (reference category 0) were associated with hazard ratios of 1.27 (95% confidence interval: 0.69, 2.53), 3.57 (95% confidence interval: 2.14, 7.48), and 6.63 (95% confidence interval: 3.57, 14.97), respectively; hazard ratios for summed segmented vessel-specific scores of 1–5, 6–11, and 12–30 (reference category 0) were 1.72 (95% confidence interval: 1.05, 3.34), 5.11 (95% confidence interval: 2.92, 10.94), and 6.10 (95% confidence interval: 3.19, 14.05), respectively; and hazard ratios for overall visual assessment of mild, moderate, or heavy (reference category none) were 2.09 (95% confidence interval: 1.30, 4.16), 3.86 (95% confidence interval: 2.02, 8.20), and 6.95 (95% confidence interval: 3.73, 15.67), respectively.

**Conclusion**—By using low-dose CT performed for lung cancer screening in older, heavy smokers, a simple visual assessment of CAC can be generated for risk assessment of CHD death and all-cause mortality, which is comparable to Agatston scoring and strongly associated with outcome.

---

Coronary artery calcification (CAC) is an established predictor of cardiovascular events and is strongly associated with advanced age and history of cigarette smoking (1). United States Preventive Services Task Force recommendations for low-dose computed tomographic (CT) screening for lung cancer, which is people aged 55–80 years and current or former smoking history of at least 30 pack-years, yield a population at risk not only for lung cancer, but also for coronary heart disease (CHD) (2). It is noteworthy that the leading cause of death in the National Lung Screening Trial (NLST) was cardiovascular illness (956 deaths) rather than lung cancer (930 deaths) (3).

CT screening for lung cancer offers an opportunity for detection of unsuspected cardiovascular disease, including CAC, and also for risk stratification for CHD events. There is not yet a consensus within the lung cancer screening community on reporting of CAC. Some investigators do not consider CAC to be clinically relevant, others consider CAC to be clinically significant only if it is extensive, whereas others suggest that CAC quantification could reduce cardiovascular morbidity and mortality and enhance the cost effectiveness of CT-based screening in heavy smokers (4–7). Our hypothesis was that a simple qualitative method of CAC scoring on low-dose CT screening for lung cancer may be sufficient to inform the patient and referring physician of the risk of cardiovascular disease in the patient.

The purpose of our study was to evaluate three CAC scoring methods to assess risk of CHD death and all-cause mortality (ACM) in NLST participants across levels of CAC scores.

## Materials and Methods

The design, eligibility criteria, participant characteristics, imaging protocols, and data collection methods of the NLST were previously described (3,8,9). Briefly, from August 2002 to April 2004, 53 454 high-risk individuals were enrolled at 33 medical centers in the United States. The NLST was approved by the institutional review board at each of the participating institutions and informed consent was obtained from all participants. Image review for this study was compliant with the Health Insurance Portability and Accountability Act. Eligibility criteria were patient age 55–74 years and 30 pack-years or more of cigarette

smoking history. Former smokers must have quit smoking within the previous 15 years. Approximately half of these participants were randomized to three rounds of annual screening with low-dose CT imaging (26 722 participants). For this analysis, participants were chosen from among the prevalence of low-dose CT scans in the American College of Radiology Imaging Network of the NLST (9413 participants). Low-dose CT imaging acquisition parameters were as follows: 120 kVp; 40–80 mAs (depending on body habitus); detector collimation, 1.25–2.5 mm; reconstruction interval, 1.0–2.5 mm; and soft-tissue reconstruction algorithm (10). Images were unenhanced and ungated.

## Study Design

We performed a retrospective, randomly selected, case-control study to analyze the relationship between baseline low-dose CT imaging of CAC and two outcomes (CHD death and ACM) that was similar to the design of the Dutch-Belgian Lung Cancer Screening Trial (Nederlands-Leuvens Longkanker Screenings onderzoek, or NELSON trial) (11). Details of the case-control sampling can be found in Figure 1. The main study included baseline (T0) low-dose CT scans from 1575 participants in the CT imaging section of the American College of Radiology Imaging Network NLST, separated into two case groups (210 CHD deaths and 315 deaths not from CHD) and one control group (1050 participants either censored or still alive at the conclusion of the trial).

Five board-certified cardiothoracic radiologists (C.C., G.W.G., J.G.R., S.G.B., and R.F.M.) participated in the study. A pilot study was first conducted on the basis of a subset of 60 CT scans, randomly selected from the main study case list, and interpreted by the five radiologists. The pilot study allowed for both reader training and the opportunity to evaluate the case report form designed to capture image interpretation. For the main study, scans were distributed evenly among the five radiologists, with radiologists blinded to participant outcomes. In addition, each reader reinterpreted the entire subset of 60 pilot CT scans during the main study to allow assessment of interreader agreement. To reduce reader recall, a washout period of a minimum of 2 months was enforced between the pilot study and the commencement of the main study.

For purposes of this analysis, death was defined from death certificate data. Specifically, CHD death was based on the following International Classification of Diseases–10 codes: I20, angina pectoris; I21, acute myocardial infarction; I22, subsequent myocardial infarction; I23, certain current complications following acute myocardial infarction; I24, other acute ischemic heart diseases; I25, chronic ischemic heart disease; I46, cardiac arrest; and I50, heart failure. Deaths in the NLST were recorded through December 31, 2009, and participants who were still alive and in follow-up were censored as of that date.

## Low-dose CT Image Analysis

Radiologists evaluated CT scan quality as adequate, suboptimal (patient or technical factors compromised image quality), or inadequate (patient or technical factors precluded assessment of CAC). CAC was assessed both quantitatively and qualitatively for all scans with adequate or suboptimal image quality. Three scoring methods were used. For scoring method 1, radiologists provided a simple, overall visual assessment of none, mild, moderate,

or heavy CAC for the entire coronary arterial circulation. For scoring method 2, radiologists performed segmented vessel-specific scoring by using an ordinal scale of 0–3. No CAC was assigned a score of 0, mild CAC was assigned a score of 1, moderate CAC was assigned a score of 2, and heavy CAC was assigned a score of 3. CAC was classified as mild if there were only isolated flecks of CAC within a segment. CAC was classified as heavy if there was continuous CAC within a segment. CAC was classified as moderate if there was more CAC than could be considered mild, but it was less than the description of heavy CAC. A set of standard images was collected to demonstrate examples of the scale (Fig 2, Fig E1 [online]). The ordinal scores from all coronary artery segments were then added together, which provided a summed segmented vessel-specific score that ranged from 0 to 30.

Coronary arteries scored included the left main, left anterior descending, left circumflex, and right coronary artery. The left anterior descending, left circumflex, and right coronary artery were each visually divided into three segments: proximal, middle, and distal. The proximal left anterior descending artery included the segment proximal to the first acute diagonal or the first 1 cm of the artery if the first acute diagonal was not visible. The middle left anterior descending artery included the segment between the first and second diagonals, or the second 1 cm of the artery if the diagonals were not visible. The distal left anterior descending artery included the remainder of the left anterior descending artery. Similarly, the left circumflex artery was divided into proximal (up to first obtuse marginal or the first 1 cm of the artery), middle (between first and second obtuse marginal or the second 1 cm), and distal (the remainder of the vessel) segments. Finally, the right coronary artery was divided into proximal (horizontal), middle (vertical), and distal (horizontal component curving under the base of the heart, which in some cases included the posterior descending coronary artery).

Scoring method 3 was a quantitative analysis of CAC, which was obtained by using a modified Agatston method with the traditional 130-HU threshold and minimal lesion definition of  $0.52 \text{ mm}^2$  (2,12,13).

### Statistical Analysis

The main goal of this analysis was to examine the association between CAC scoring, measured on NLST low-dose CT scans, and both CHD death and ACM. Risk categories for scoring method 1 (the overall visual assessment) included none, mild, moderate, and heavy calcification. For scoring method 2 (summed segmented vessel-specific scoring), the risk categories included 0, 1–5, 6–11, and 12–30. Risk categories for scoring method 3 (Agatston scoring) included 0, 1–100, 101–1000, and greater than 1000.

Interreader agreement was assessed by using the intraclass correlation coefficient for continuous measures (scoring methods 2 and 3) and the weighted  $\kappa$  statistic for ordinal measures (scoring method 1). For each statistic, agreement was summarized by the mean and range over all 10 pairwise combinations of the five radiologists. The following intraclass correlation coefficient scale, proposed by Meyers (14), was used to determine the extent of

---

*Online supplemental material is available for this article.*

agreement: poor (<0.7), fair (between 0.7 and 0.8), good (between 0.8 and 0.9), and high (0.9).

Baseline characteristics were first summarized and compared across the three sampling groups. For continuous variables, one-way analysis of variance was used; for categorical variables,  $\chi^2$  statistics were computed. Next, for each scoring method, event rates for both CHD death and ACM were reported. Event rates were calculated by dividing the number of identified cases in each CAC category (ie, the numerator) by the predicted total number of cases in that category among all American College of Radiology Imaging Network low-dose CT imaging participants with a baseline T0 scan and coded death certificate data (ie, the denominator). Because the sampling selection weights were needed to predict the total number of cases, we refer to this rate as the selection-weighted absolute event rate.

Finally, weighted Cox proportional hazards regression was used to estimate the association between CAC and outcomes, summarized with hazard ratios and associated 95% confidence intervals (11). Because the data were derived from a retrospective case control study, a weighted model was used to provide unbiased estimates of the hazard ratios, where the weights for each sampling group corresponded to the inverse of their respective sampling probabilities (15). The bootstrap technique was applied to derive 95% confidence intervals for all reported hazard ratios. Separate models were fit for time to CHD death and time to ACM. In the former, deaths not from CHD were censored at the time of death. In the latter, both CHD and deaths not from CHD were treated as events. For each outcome, both univariate and multivariate models were fit. In the multivariate models, we adjusted for potential confounders (defined by patient-completed questionnaires at baseline), including age, sex, race, ethnicity, education, marital status, ever diagnosed with diabetes, ever diagnosed with hypertension, chronic obstructive pulmonary disease, and smoking status. A history of diabetes, hypertension, and chronic obstructive pulmonary disease were defined by yes or no answers to the following question: "Has a doctor ever told you that you have the following conditions or illnesses?"

Although unconditional logistic regression might be a natural method to analyze data from a case-control study, Cox proportional hazards regression can be used after implementation of a proper weighting mechanism to accommodate the retrospective design of the study (15). We chose the latter approach to compare our results to those of the NELSON study (11).

## Results

There was evidence of previous revascularization or stent on 86 of 1575 CT scans (5.5%), and these were excluded from further analysis (Fig 1). Among the remaining CT scans, image quality was assessed as adequate in 93.4% (1390 of 1489 CT scans), as suboptimal in 3.5% (52 of 1489 CT scans), and as inadequate in 3.2% (47 of 1489 CT scans). CT scans were considered inadequate because of incomplete anatomic coverage (four CT scans), unacceptable image noise (36 CT scans), or other reasons (seven CT scans). Results are based on the 1442 scans with adequate or suboptimal image quality. The numbers of CT scans interpreted by each reader are as follows: 287 CT scans were interpreted by C.C., 277 CT scans were interpreted by G.W.G., 290 CT scans were interpreted by J.G.R., 296 CT

scans were interpreted by R.F.M., and 292 CT scans were interpreted by S.G.B. These included 171 CHD deaths, 295 deaths not from CHD, and 976 control participants. Median follow-up time for the control group was 2304 days (interquartile range, 2213–2441 days), with 16 of 976 participants (1.6%) censored before December 31, 2009.

Baseline characteristics of the study population and the entire NLST cohort are shown in Table 1. Within the CHD death group, compared with the control group, participants were more likely to be older ( $P < .001$ ), men ( $P = .002$ ), current smokers ( $P < .001$ ), and previously diagnosed with diabetes ( $P < .001$ ) and hypertension ( $P < .001$ ). Within the group of deaths that were not from CHD, compared with the control group, participants were more likely to be older ( $P < .001$ ), men ( $P < .001$ ), not married ( $P < .001$ ), less educated ( $P < .001$ ), current smokers ( $P < .001$ ), and previously diagnosed with diabetes ( $P < .001$ ).

CAC scores corresponding to each scoring method are summarized in Table 2. A CAC score corresponding to no calcification was assigned to 387 CT scans (26.8%) by using scoring method 1 (overall visual assessment), 383 CT scans (26.6%) by using scoring method 2 (summed segmented vessel-specific scoring), and 400 CT scans (27.7%) by using scoring method 3 (Agatston scoring). An Agatston score of 0 was present in 33.5% (327 of 976) of control participants, 11.7% (20 of 171) of patients with CHD death, and 18.0% (53 of 295) of patients with death not from CHD ( $P < .001$ ).

The overall visual assessment exhibited good agreement with the categorized Agatston scores (weighted  $\kappa$ , 0.75 [95% confidence interval: 0.73, 0.77]), with radiologists assigning participants to the same risk category as the Agatston score in 73.0% (1052 of 1442) of CT scans and to within one category in 99.7% (1438 of 1442) of CT scans. The interreader agreement was good or high for all three scoring methods: mean weighted  $\kappa$  of 0.85 (range, 0.74–0.95) for scoring method 1, mean intraclass correlation coefficient of 0.88 (range, 0.78–0.97) for scoring method 2, mean intraclass correlation coefficient of 0.92 (range, 0.84–1.00) for scoring method 3.

Selection-weighted absolute event rates and hazard ratios for time to CHD death and ACM are shown in Tables 3 and 4. The simplest CAC evaluation, scoring method 1, was strongly associated with both outcomes and demonstrated increased risk of event with increased level of calcification, even after we adjusted for potential confounders.

## Discussion

To our knowledge, there is not yet a consensus on whether to include CAC as a significant incidental finding on low-dose CT performed for lung cancer screening or how to report CAC. Although the Agatston scoring method is well validated for measurement of CAC on gated CT imaging, we demonstrated that CAC evaluation on ungated low-dose CT imaging originally performed for indications other than calcium scoring is adequate for cardiac risk stratification. This may eliminate the need for an additional, dedicated calcium scoring CT in this population with increased risk of CHD. The simplest method, an overall visual assessment of CAC as none, mild, moderate, or heavy, was able to separate patients into risk categories on the basis of either CHD death or ACM. All five radiologists in this study



preferred the visual assessment, not only because it was faster and simpler than either of the other two methods, but it also eliminated the need for additional software.

These findings confirm and extend similar findings on the basis of other lung cancer screening studies. We confirm that Agatston scoring and ordinal scoring methods are predictive of cardiovascular death and ACM, but also demonstrate that a simpler so-called Gestalt method of visual analysis may be sufficient for risk classification. Jacobs et al (11) reported that CAC scoring with Agatston scores from low-dose CT scans in the NELSON study could be used as an independent predictor of cardiovascular events and ACM. In a retrospective case-control study of 150 cases and 808 control participants, Jacobs et al reported hazard ratios similar to our study for prediction of cardiovascular events and deaths. In a subsequent analysis of NELSON trial participants, Mets et al (16) used automated quantification of coronary and aortic calcium volume with patient age, cardiovascular history, and smoking characteristics to derive and validate a prediction model for cardiovascular events. Similarly, based on the Multicentric Italian Lung Detection project, Sverzellati et al (17) reported that an Agatston score greater than 400 was independently associated with both cardiovascular events and ACM.

Slightly over 10% of the subjects in the CHD-death group in our study had no visible CAC. In an evaluation of the concordance of CAC measurement in 483 patients with both low-dose ungated and regular-dose electrocardiogram-gated multidetector-row CT imaging, Wu et al (13) noted that two observers had four (0.8%) and five (1.0%) false-negative findings by using low-dose CT imaging, which corresponded with Agatston scores of 1–12 on gated studies. Although our finding that 10% of CHD deaths occurred in patients with no calcification at ungated low-dose CT imaging may overestimate this risk, it does support the statement by McEvoy et al (18) that the absence of CAC might not be as useful as a so-called negative risk factor in active smokers compared with nonsmokers. Although the absence of CAC in asymptomatic patients predicts excellent survival, with 10-year event rates of approximately 1%, mortality is higher in smokers (ACM of 3.31 deaths per 1000 person-years) with CAC scores of 0 compared with nonsmokers (0.67 deaths per 1000 person-years) (18,19). This fact serves to highlight that the older, smoking population eligible for low-dose CT lung cancer screening may be unaware of their increased risk of coronary artery disease. Age and smoking history alone may be useful to predict the risk of coronary atherosclerosis and to motivate patients to make lifestyle changes and initiate or adhere to lipid-lowering therapy.

Techniques for low-dose CT screening for lung cancer are unlike those for cardiac CT screening because low-dose CT scans are ungated and have a lower signal-to-noise ratio. Nevertheless, CAC is identifiable and measurable. Budoff et al (20) compared Agatston CAC scores on gated and ungated CT scans obtained in 50 participants in the chronic obstructive pulmonary disease gene trial. Although Agatston scores obtained with ungated CT scans tended to be higher than those obtained with gated CT scans (respective mean absolute Agatston values, 353.6 vs 277.1), there was excellent correlation for stratification into risk categories of 0, 1–100, 101–400, and greater than 400.

Similar to our study, other investigators (21–24) used semiquantitative visual assessments of CAC by using CT scans not obtained with technical parameters of cardiac CT and have demonstrated association with outcomes. Einstein et al (22) reported a high degree of association between visual estimation of CAC and Agatston scoring on ungated low-dose CT scans used for attenuation correction on positron emission tomography/CT and single photon emission CT/CT. Visual estimates of CAC on a six-level scale were within the same category as the Agatston score category in 63% of cases and within one category in 93%. Shemesh et al (25) used an ordinal scoring technique to evaluate the frequency of CAC in 4250 participants in the Early Lung Cancer Action Program. In a population of current and former smokers, CAC was present in 64% of participants. Correlation of CAC scores with cardiovascular death, in a cohort of 8782 smokers aged 40–85 years (23) with a median follow-up of 6 years, revealed 193 deaths from cardiovascular disease. The hazard ratio for a CAC score greater than 4, adjusted for age, sex, pack-years of smoking, and diabetes, was 2.1, compared with the reference category of a CAC score of 0.

There are some limitations to our study. In the absence of NLST data regarding cholesterol levels, systolic blood pressure, and antihypertensive medication, we could not calculate Framingham risk scores, and so we could not determine if the CAC score further improves risk stratification in this lung cancer screening population. An additional limitation is that CHD death was determined solely from the death certificate, without further verification. We also note that the heterogeneity of low-dose CT imaging acquisition parameters within the NLST, with variable image thicknesses and reconstruction intervals, may influence both visual and Agatston scores.

CAC scoring is presently not a component of low-dose CT screening for lung cancer. On the basis of our results, we encourage including the presence of CAC as a clinically significant finding on low-dose CT screening for lung cancer. This can be reported as a modified Agatston score with the caveat that the number is likely to be higher than a dedicated calcium scoring CT scan would demonstrate. However, we feel that our simplest analysis, the overall visual assessment, is comparable to Agatston scoring and could be easily incorporated into structured reporting systems for lung cancer screening, such as the Lung Imaging Reporting and Data System, and might be more readily adopted into mass screening programs. This would allow for built-in follow-up recommendations, which would therefore standardize the process and reduce its overuse.

On ungated low-dose CT scans, originally performed for purposes of lung cancer screening, a simple overall visual assessment of CAC as none, mild, moderate, or heavy can separate patients into risk categories on the basis of either CHD death or ACM, and demonstrates good interreader agreement.

## Abbreviations

<b>ACM</b>	all-cause mortality
<b>CAC</b>	coronary artery calcification
<b>CHD</b>	coronary heart disease



## NLST National Lung Screening Trial

### References

1. Greenland P, Bonow RO, Brundage BH, et al. ACCF/AHA 2007 clinical expert consensus document on coronary artery calcium scoring by computed tomography in global cardiovascular risk assessment and in evaluation of patients with chest pain: a report of the American College of Cardiology Foundation Clinical Expert Consensus Task Force (ACCF/AHA Writing Committee to Update the 2000 Expert Consensus Document on Electron Beam Computed Tomography) developed in collaboration with the Society of Atherosclerosis Imaging and Prevention and the Society of Cardiovascular Computed Tomography. *J Am Coll Cardiol.* 2007; 49(3):378–402. [PubMed: 17239724]
2. Moyer VA. U.S. Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2014; 160(5):330–338. [PubMed: 24378917]
3. Aberle DR, Adams AM, et al. National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011; 365(5):395–409. [PubMed: 21714641]
4. Kucharczyk MJ, Menezes RJ, McGregor A, Paul NS, Roberts HC. Assessing the impact of incidental findings in a lung cancer screening study by using low-dose computed tomography. *Can Assoc Radiol J.* 2011; 62(2):141–145. [PubMed: 20382501]
5. Priola AM, Priola SM, Gajj-Levra M, et al. Clinical implications and added costs of incidental findings in an early detection study of lung cancer by using low-dose spiral computed tomography. *Clin Lung Cancer.* 2013; 14(2):139–148. [PubMed: 22868220]
6. van de Wiel JC, Wang Y, Xu DM, et al. Neglectable benefit of searching for incidental findings in the Dutch-Belgian lung cancer screening trial (NELSON) using low-dose multidetector CT. *Eur Radiol.* 2007; 17(6):1474–1482. [PubMed: 17206426]
7. Mets OM, de Jong PA, Prokop M. Computed tomographic screening for lung cancer: an opportunity to evaluate other diseases. *JAMA.* 2012; 308(14):1433–1434. [PubMed: 23047354]
8. Aberle DR, Berg CD, et al. National Lung Screening Trial Research Team. The National Lung Screening Trial: overview and study design. *Radiology.* 2011; 258(1):243–253. [PubMed: 21045183]
9. Aberle DR, Adams AM, et al. National Lung Screening Trial Research Team. Baseline characteristics of participants in the randomized national lung screening trial. *J Natl Cancer Inst.* 2010; 102(23):1771–1779. [PubMed: 21119104]
10. Cagnon CH, Cody DD, McNitt-Gray MF, Seibert JA, Judy PF, Aberle DR. Description and implementation of a quality control program in an imaging-based clinical trial. *Acad Radiol.* 2006; 13(11):1431–1441. [PubMed: 17111584]
11. Jacobs PC, Gondrie MJ, van der Graaf Y, et al. Coronary artery calcium can predict all-cause mortality and cardiovascular events on low-dose CT screening for lung cancer. *AJR Am J Roentgenol.* 2012; 198(3):505–511. [PubMed: 22357989]
12. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M Jr, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol.* 1990; 15(4): 827–832. [PubMed: 2407762]
13. Wu MT, Yang P, Huang YL, et al. Coronary arterial calcification on low-dose ungated MDCT for lung cancer screening: concordance study with dedicated cardiac CT. *AJR Am J Roentgenol.* 2008; 190(4):923–928. [PubMed: 18356438]
14. Meyers, CR. Measurement in physical education. New York, NY: Ronald; 1974.
15. Thernau, TM.; Grambsch, PM. Modeling survival data: extending the Cox model. New York, NY: Springer; 2000.
16. Mets OM, Vliegenthart R, Gondrie MJ, et al. Lung cancer screening CT-based prediction of cardiovascular events. *JACC Cardiovasc Imaging.* 2013; 6(8):899–907. [PubMed: 23769488]

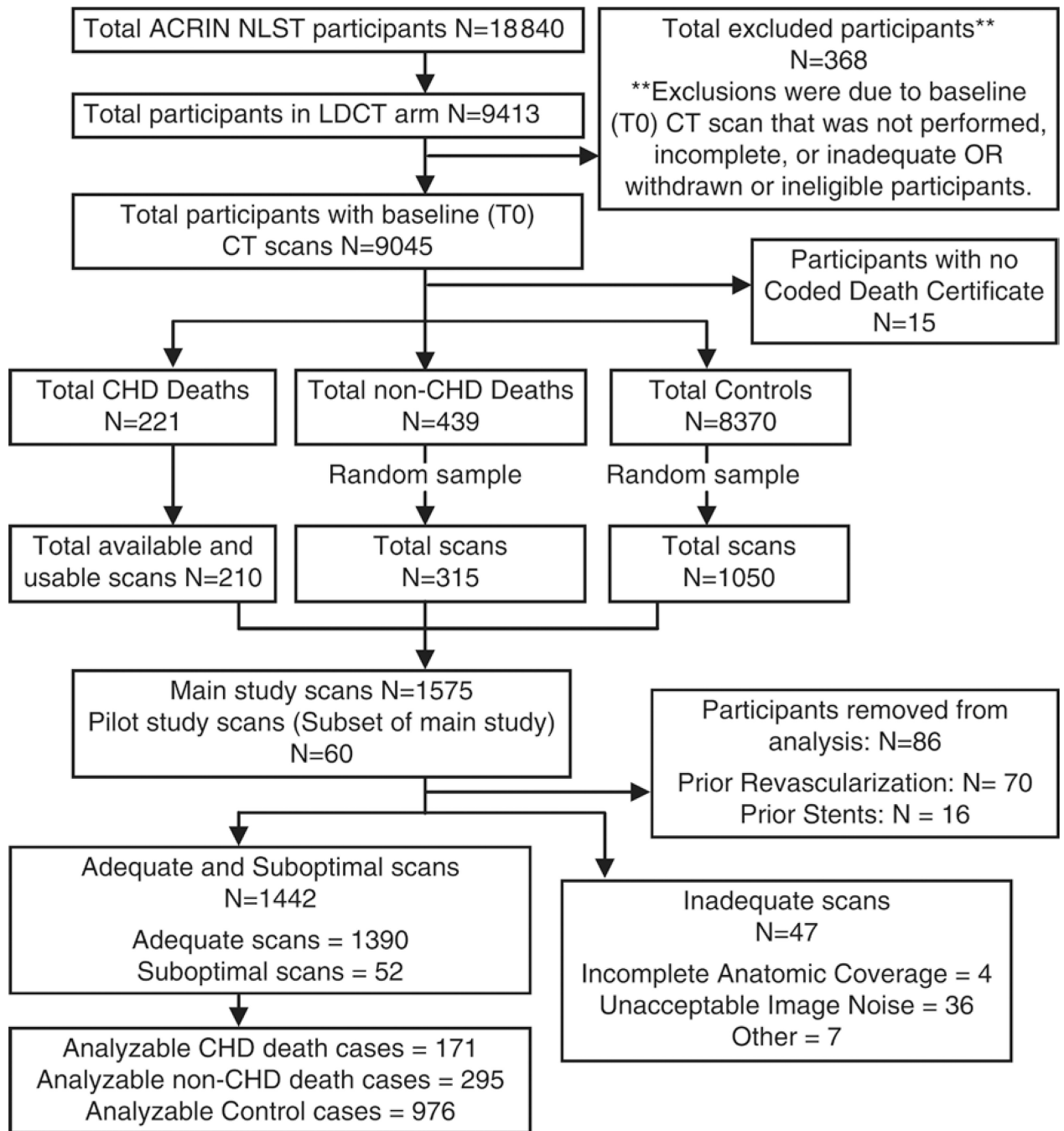
17. Sverzellati N, Cademartiri F, Bravi F, et al. Relationship and prognostic value of modified coronary artery calcium score, FEV1, and emphysema in lung cancer screening population: the MILD trial. *Radiology*. 2012; 262(2):460–467. [PubMed: 22114241]
18. McEvoy JW, Blaha MJ, Rivera JJ, et al. Mortality rates in smokers and nonsmokers in the presence or absence of coronary artery calcification. *JACC Cardiovasc Imaging*. 2012; 5(10):1037–1045. [PubMed: 23058072]
19. Blaha M, Budoff MJ, Shaw LJ, et al. Absence of coronary artery calcification and all-cause mortality. *JACC Cardiovasc Imaging*. 2009; 2(6):692–700. [PubMed: 19520338]
20. Budoff MJ, Nasir K, Kinney GL, et al. Coronary artery and thoracic calcium on noncontrast thoracic CT scans: comparison of ungated and gated examinations in patients from the COPD Gene cohort. *J Cardiovasc Comput Tomogr*. 2011; 5(2):113–118. [PubMed: 21167806]
21. Kirsch J, Buitrago I, Mohammed TL, Gao T, Asher CR, Novaro GM. Detection of coronary calcium during standard chest computed tomography correlates with multidetector computed tomography coronary artery calcium score. *Int J Cardiovasc Imaging*. 2012; 28(5):1249–1256. [PubMed: 21833776]
22. Einstein AJ, Johnson LL, Bokhari S, et al. 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. *J Am Coll Cardiol*. 2010; 56(23):1914–1921. [PubMed: 21109114]
23. Shemesh J, Henschke CI, Shaham D, et al. Ordinal scoring of coronary artery calcifications on low-dose CT scans of the chest is predictive of death from cardiovascular disease. *Radiology*. 2010; 257(2):541–548. [PubMed: 20829542]
24. Huang YL, Wu FZ, Wang YC, et al. Reliable categorisation of visual scoring of coronary artery calcification on low-dose CT for lung cancer screening: validation with the standard Agatston score. *Eur Radiol*. 2013; 23(5):1226–1233. [PubMed: 23239060]
25. Shemesh J, Henschke CI, Farooqi A, et al. Frequency of coronary artery calcification on low-dose computed tomography screening for lung cancer. *Clin Imaging*. 2006; 30(3):181–185. [PubMed: 16632153]

### Advances in Knowledge

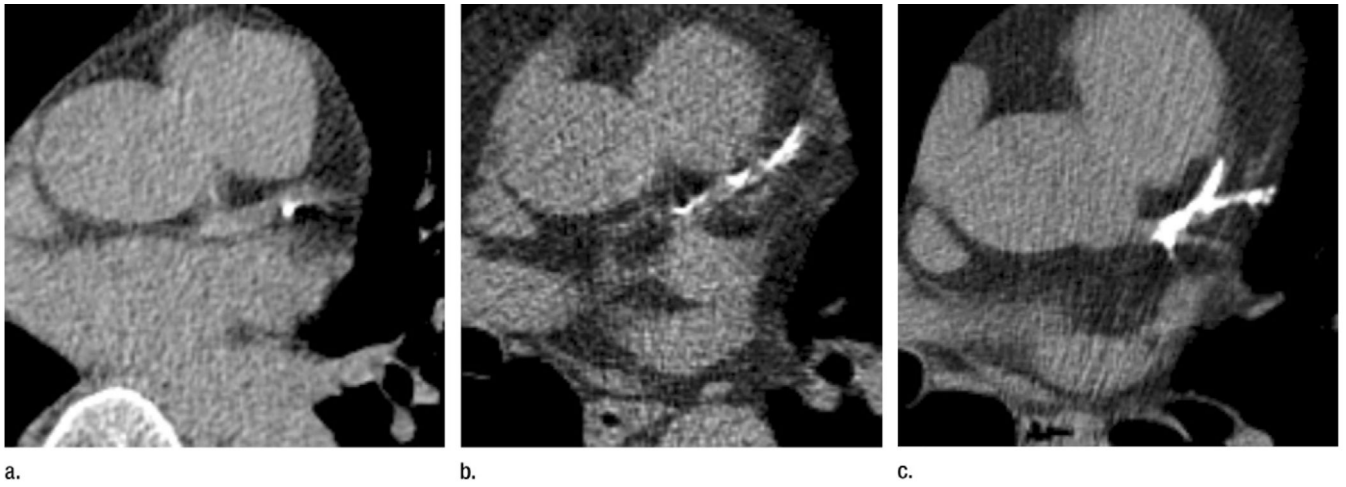
- A proposed simple method of coronary artery calcification (CAC) scoring on the basis of ungated low-dose CT scans (ie, an overall visual assessment of none, mild, moderate, or heavy) separated patients into risk categories of either coronary heart disease death or all-cause mortality.
- The overall visual assessment of CAC exhibited fair agreement with categorized Agatston scores (weighted  $\kappa = 0.75$  [95% confidence interval: 0.73, 0.77]); radiologists assigned participants to the same risk category as the Agatston score in 73.0% (1052 of 1442) of scans and to within one category in 99.7% (1438 of 1442) of scans.
- The overall visual assessment of CAC also demonstrated good interreader agreement among different radiologists (mean weighted  $\kappa = 0.85$  [range, 0.74–0.95]).

### Implications for Patient Care

- CAC can be reported as a modified Agatston score, but the number is likely to be higher than a dedicated calcium-scoring CT examination would demonstrate.
- Our simplest analysis, an overall visual assessment of CAC as none, mild, moderate, or heavy, is comparable to Agatston scoring and can be easily incorporated into structured reporting systems for lung cancer screening.



**Figure 1.** Analysis flowchart. *ACRIN*= American College of Radiology Imaging Network, *LDCT*= low-dose CT.



**Figure 2.**

(a) An example of mild CAC (score of 1). (b) An example of moderate CAC (score of 2). (c) An example of heavy CAC (score of 3).



**Table 1**  
Baseline Characteristics of the Analyzed Participants, Overall and by Sampling Group

Parameter	NLST Cohort (n = 53 452)*	Sampling Group			P Value†
		Analysis Population (n = 1442)	CHD Deaths (n = 171)	Deaths not from CHD (n = 976)	
Mean age (y)‡	61 ± 5	62 ± 5	63 ± 5	61 ± 5	<.001
Men	31 530 (59)	846 (59)	115 (67)	532 (55)	<.001
Race					.099
White	48 549 (91)	1331 (92)	157 (92)	263 (89)	...
Black	2376 (4)	90 (6)	10 (6)	28 (9)	...
Other	2527 (5)	21 (1)	4 (2)	4 (1)	...
Education level					<.001
<12th grade	3249 (6)	104 (7)	18 (11)	30 (10)	...
High school graduate or general educational development test	12 712 (24)	336 (23)	33 (19)	87 (29)	...
Training after high school	7434 (14)	162 (11)	23 (13)	23 (8)	...
Associate degree or some college	12 277 (23)	314 (22)	41 (24)	68 (23)	...
Bachelor degree	8946 (17)	253 (18)	25 (15)	37 (13)	...
Graduate school	7600 (14)	218 (15)	25 (15)	33 (11)	...
Unknown/Other	1234 (2)	55 (4)	6 (4)	17 (6)	...
Married	35 589 (67)	887 (62)	101 (59)	158 (54)	.003
Diabetes§	5174 (10)	138 (10)	32 (19)	39 (13)	<.001
Hypertension¶	18 930 (35)	508 (35)	82 (48)	111 (38)	<.001
Current smoker	25 760 (48)	773 (54)	109 (64)	190 (64)	<.001

Note.—Data are number of participants and data in parentheses are percentages unless otherwise noted. Baseline characteristics of the entire NLST cohort are added for reference.

\* Number of participants less than that originally reported in the primary NLST manuscript because of two duplicate registrations discovered after publication.

† P-values correspond to comparisons of the distribution of each variable among the three sampling groups.

‡ Data are ± standard deviation.

§ Self-reported diagnosis of diabetes.  
// Self-reported diagnosis of hypertension.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Distribution of CAC Scores for Each Scoring Method, Stratified by Sampling Group

Parameter	Sampling Group			Total (n = 1442)
	CHD Deaths (n = 171)	Deaths not from CHD (n = 295)	Control Participants (n = 976)	
CAC scoring method 1*				
None	18 (10.5)	52 (17.6)	317 (32.5)	387 (26.8)
Mild	64 (37.4)	120 (40.7)	444 (45.5)	628 (43.6)
Moderate	40 (23.4)	62 (21.0)	127 (13.0)	229 (15.9)
Heavy	49 (28.7)	61 (20.7)	88 (9.0)	198 (13.7)
CAC scoring method 2*	8.18 ± 6.92	6.48 ± 6.46	3.67 ± 4.92	4.78 ± 5.77
0	18 (10.5)	50 (16.9)	315 (32.3)	383 (26.6)
1–5	52 (30.4)	120 (40.7)	436 (44.7)	608 (42.2)
6–11	58 (33.9)	61 (20.7)	138 (14.1)	257 (17.8)
12–30	43 (25.1)	64 (21.7)	87 (8.9)	194 (13.5)
CAC scoring method 3*	720.31 ± 971.33	540.06 ± 881.99	248.50 ± 544.12	364.09 ± 707.66
0	20 (11.7)	53 (18.0)	327 (33.5)	400 (27.7)
1–100	27 (15.8)	67 (22.7)	295 (30.2)	389 (27.0)
101–1000	82 (48.0)	127 (43.1)	284 (29.1)	493 (34.2)
>1000	42 (24.6)	48 (16.3)	70 (7.2)	160 (11.1)

Note.—Data are number of participants and data in parentheses are percentages, except where indicated.  $P < .001$  between the three sampling groups.

\*Data are mean ± standard deviation.

**Table 3**

Association between CAC and Time to CHD Death for Each of Three Scoring Methods

CAC Category	CAC Scoring Method 1	CAC Scoring Method 2	CAC Scoring Method 3
Scoring method 1 score of none, scoring method 2 score of 0, or scoring method 3 score of 0			
Selection-weighted absolute event rates for CHD death*	18/2819 (0.6)	18/2799 (0.6)	20/2909 (0.7)
Model-based hazard ratios			
Univariate model	1	1	1
Multivariate model	1	1	1
Scoring method 1 score of mild, scoring method 2 score of 1–5, or scoring method 3 score of 1–100			
Selection-weighted absolute event rates for CHD death*	64/4069 (1.6)	52/3985 (1.3)	27/2664 (1.0)
Model-based hazard ratios			
Univariate model	2.48 (1.55, 4.55)	2.03 (1.27, 3.66)	1.46 (0.84, 2.66)
Multivariate model	2.09 (1.30, 4.16)	1.72 (1.05, 3.34)	1.27 (0.69, 2.53)
Scoring method 1 score of moderate, scoring method 2 score of 6–11, or scoring method 3 score of 101–1000			
Selection-weighted absolute event rates for CHD death*	40/1233 (3.2)	58/1349 (4.3)	82/2731 (3.0)
Model-based hazard ratios			
Univariate model	5.18 (2.97, 9.69)	6.88 (4.15, 12.75)	4.53 (2.96, 8.13)
Multivariate model	3.86 (2.02, 8.20)	5.11 (2.92, 10.94)	3.57 (2.14, 7.48)
Scoring method 1 score of heavy, scoring method 2 score of 12–30, or scoring method 3 score of >1000			
Selection-weighted absolute event rates for CHD death*	49/909 (5.4)	43/897 (4.8)	42/726 (5.8)
Model-based hazard ratios			
Univariate model	9.04 (5.56, 16.87)	8.11 (4.85, 15.19)	9.11 (5.34, 16.76)
Multivariate model	6.95 (3.73, 15.67)	6.10 (3.19, 14.05)	6.63 (3.57, 14.97)

Note.—Data in parentheses are 95% confidence interval ranges unless otherwise indicated. The multivariate analysis was adjusted for age, sex, race, ethnicity, education (baseline), marital status (baseline), ever diagnosed with diabetes (baseline), ever diagnosed with hypertension (baseline), chronic obstructive pulmonary disease status (baseline), and smoking status (baseline).

\* Data in parentheses are percentages. For each reported selection-weighted absolute event rate, the numerator corresponds to the number of identified events in the respective CAC category and the denominator corresponds to the selection-weighted total number of participants predicted by the scoring method among all American College of Radiology Imaging Network low-dose CT imaging participants with a baseline T0 scan and coded death certificate data ( $n = 9030$ ; Fig 1).

**Table 4**

Association between CAC and Time to ACM for Each of Three Scoring Methods

CAC Category	CAC Scoring Method 1	CAC Scoring Method 2	CAC Scoring Method 3
Scoring method 1 score of none, scoring method 2 score of 0, or scoring method 3 score of 0			
Selection-weighted absolute event rates for ACM*	70/2819 (2.5)	68/2799 (2.4)	73/2909 (2.5)
Model-based hazard ratios			
Univariate model	1	1	1
Multivariate model	1	1	1
Scoring method 1 score of mild, scoring method 2 score of 1–5, or scoring method 3 score of 1–100			
Selection-weighted absolute event rates for ACM*	184/4069 (4.5)	172/3985 (4.3)	94/2664 (3.5)
Model-based hazard ratios			
Univariate model	1.79 (1.30, 2.44)	1.77 (1.29, 2.45)	1.39 (0.97, 1.94)
Multivariate model	1.50 (1.06, 2.11)	1.48 (1.05, 2.14)	1.20 (0.82, 1.73)
Scoring method 1 score of moderate, scoring method 2 score of 6–11, or scoring method 3 score of 101–1000			
Selection-weighted absolute event rates for ACM*	102/1233 (8.3)	119/1349 (8.8)	209/2731 (7.7)
Model-based hazard ratios			
Univariate model	3.29 (2.27, 4.74)	3.54 (2.50, 5.09)	3.07 (2.25, 4.22)
Multivariate model	2.39 (1.60, 3.65)	2.49 (1.69, 3.76)	2.30 (1.67, 3.38)
Scoring method 1 score of heavy, scoring method 2 score of 12–30, or scoring method 3 score of >1000			
Selection-weighted absolute event rates for ACM*	110/909 (12.1)	107/897 (11.9)	90/726 (12.4)
Model-based hazard ratios			
Univariate model	4.99 (3.59, 7.27)	5.16 (3.57, 7.58)	5.10 (3.58, 7.61)
Multivariate model	3.46 (2.33, 5.54)	3.65 (2.44, 5.81)	3.52 (2.40, 5.69)

Note.—Data in parentheses are 95% confidence interval ranges unless otherwise indicated. The multivariate analysis was adjusted for age, sex, race, ethnicity, education (baseline), marital status (baseline), ever diagnosed with diabetes (baseline), ever diagnosed with hypertension (baseline), chronic obstructive pulmonary disease status (baseline), and smoking status (baseline).

\* Data in parentheses are percentages. For each reported selection-weighted absolute event rate, the numerator corresponds to the number of identified events in the respective CAC category and the denominator corresponds to the selection-weighted total number of participants predicted by the scoring method among all American College of Radiology Imaging Network low-dose CT imaging participants with a baseline T0 scan and coded death certificate data ( $n = 9030$ ; Fig 1).